



Recommendation System for LastFM

11 March 2022

Hina Hussain

Inderpreet Rana

Mui Han Ma

Recommendation Tools by Philipp Borchert

A. Introduction

LastFM is a well-known music streaming service which brings together your favorite music services and join up listening, watching and sharing to connect your musical world. The objective of this report is to help LastFM improve their recommendation system to provide their customers with a better experience and increase the amount time spent on the platform. Currently LastFM is using a **non-personalized** recommender system by recommending the 10 most popular artists to all users.

This report runs multiple benchmark recommendation models such as **collaborative filtering** algorithms, **content-based** algorithms, and **hybrid** algorithms. The best performing model is selected using various accuracy measures and a **personalized** recommendation system is proposed using it.

B. Collaborative Filtering Recommendation Systems

Collaborative filtering (CF) is a technique that identifies items that a user might like on the basis of reactions by similar users. It searches all the users and identifies a smaller set of users with tastes similar to a particular user. It also looks at the items they like and combines them to create a ranked list of suggestions. There are many ways to measure similarity but the main are cosine and pearson similarity.

For this project, the following five CF algorithms were run using **default** parameters in the surprise package. The algorithms were run using **cross validation** to train and test. Cross validation is a resampling technique for assessing how the statistical analysis generalises to an independent data set. It is used to evaluate the algorithms by training multiple models on subsets of the dataset and then evaluates them on another subset of the data. The project used the cross_validate function of the surprise package with the cv iterator set to 3. This function uses k-fold cross validation. K-fold splits the data into k subsets and ensures that every subset from the original dataset has the chance of appearing in training and test set. The whole user-item-rating dataset is used in this section.

1. K Nearest Neighbors – KNNBasic

KNNBasic is a basic collaborative algorithm. KNN approach can either be item-based or user-based. KNN relies on using item feature similarity. It calculates the distance between the target item/user and then it ranks the distance and returns the top K nearest neighbor items/users as the most similar recommendations. The KNNbasic algorithm uses the max number of neighbors to use for aggregation, the minimum number of neighbors to use for aggregation, and the similarity measure (either Cosine or Pearson measure) to calculate the predictions¹.

Advantages

- Simple and easy – this algorithm is easy to understand and implement
- Memory based approach – there is no assumptions & training step, the model learns by using historical data

¹ Hina Hussain; Individual Assignment for Recommendation Tools

Disadvantages

- Sensitive to outliers – as this model learns by using historical data, outliers (if present) are also included and put in the existing groups leading to stronger deviation in the model
- Difficult to choose the optimal number of neighbors – if the K is not chosen correctly, the number of groups is not correct and hence the model can be underfit or overfit.

2. Singular Value Decomposition (SVD)

SVD is a matrix factorization technique. It uses a user-item matrix and then decomposes the matrix using SVD into 3 matrices. The three matrices represent the relationship between users and high-level latent factors, the strength of each latent factor, and the similarity between items and latent factors. The main concept is to decompose the original and very sparse matrix into two low-rank matrices that represent user factors and item factors. Stochastic Gradient Descent is used to minimize the loss function in question.

The SVD function of the surprise package was used with default parameters. The main parameters are the number of factors, number of epochs which is the number of times the minimization step (minimize the squared error) is performed on the train set, whether to use a biased or unbiased version of the algorithm, the learning rate of all the parameters and the regularization term for all parameters².

Advantages

- Efficient as it can be applied to a matrix with multiple features
- Precise result for the ranking as it transforms the features into matrix and consider them individually

Disadvantages

- Does not work well for non-linear data
- Difficult to interpret the matrix

3. SVD++

This algorithm is an extension of the SVD and takes into account the implicit ratings of users. Implicit means that chances that the user "likes" an item he/she has rated are higher than for a random not-rated item.

The parameters used in the SVD++ algorithm are the same as SVD with the exception of bias.

Advantages

- More accurate result as compared to SVD. Since it takes implicit ratings into account, it has more user relevant information to train the model

² Hina Hussain; Individual Assignment for Recommendation Tools

Disadvantage

- Longer computation time. In this case, it also has to consider whether the user listened to that artist or not

4. Baseline Only

This algorithm transforms the rating matrix into user matrix (i.e. platform users) and item matrix (i.e. artists) which is the baseline. The baseline is built by either using Alternating Least Squares (ALS) or Stochastic Gradient Descent (SGD) method.

Advantages

- Easy to implement
- Fast computation time

Disadvantage

- As it is calculated by transforming the rating matrix into user matrix and item matrix, there is not many parameters that can be tuned to optimize the model compared to other models

5. Co-Clustering

This approach assigns users and items with their own clusters and also co-clusters. Clusters are assigned using a straightforward optimization method, similar to k-means. The predictions are made using the average ratings of the cluster particular user, average rating of the cluster particular item, and the average rating of the co-cluster of that item and user.

The CoClusteting function of the surprise package was used with default parameters for this project. The main parameters are number of user clusters, number of item clusters, and the number of iterations of the optimization loop³.

Advantages

- Lower computation resource used

Disadvantages

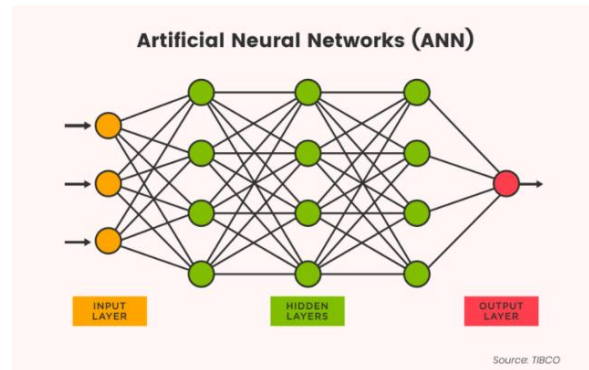
- Complex to interpret

6. Neural Networks

Artificial neural networks (ANNs) and simulated neural networks (SNNs) are a subset of machine learning that are at the heart of deep learning methods. Their name and structure are derived from the human brain, and they resemble the way biological neurons communicate with one another.

³ Hina Hussain; Individual Assignment for Recommendation Tools

Artificial neural networks (ANNs) are comprised of a node layer, containing an input layer, one or more hidden layers, and an output layer. Each node, or artificial neuron, is connected to the others and has a weight and threshold linked with it. If a node's output exceeds a certain threshold value, the node is activated, and data is sent to the next tier of the network. Otherwise, no data is sent on to the network's next tier.



Think of each individual node as its own linear regression model, composed of input data, weights, a bias (or threshold), and an output. The formula would look something like this:

$$\sum w_i x_i + \text{bias} = w_1 x_1 + w_2 x_2 + w_3 x_3 + \text{bias}$$

$$\text{output} = f(x) = 1 \text{ if } \sum w_i x_i + b \geq 0; 0 \text{ if } \sum w_i x_i + b < 0$$

Recommendation systems in neural networks:

Real-world neural network recommender systems are often composed of two stages:

The **retrieval stage** oversees selecting a small number of applicants from a large pool of candidates. The main objective of this model is to efficiently weed out all candidates that the user is not interested in. Because the retrieval model may be dealing with millions of candidates, it must be computationally efficient.

In the 2nd stage, that is the **ranking stage** the retrieval model's outputs are fine-tuned in the ranking stage to select the finest feasible handful of suggestions. Its goal is to reduce the number of items the user could be interested into a manageable number.

In this project we only focused on the retrieval stage.

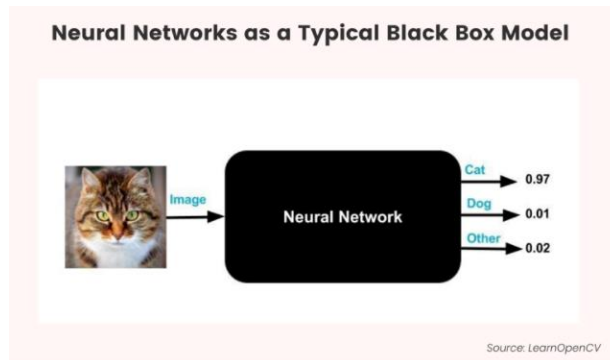
Advantages

- The ability to learn by themselves. As the base of profound learning, neural systems can perform unsupervised learning and can deliver yields that are not constrained to the input given to them

- The ability to work with insufficient data and information. Even if the data is incomplete or insufficient, the network can detect the error and still produce the output. That's because the entire output generation is not affected by the corruption of one or more than one neuron.
- The capacity of parallel preparing. Neural systems are able of performing numerous errands at a time without influencing the framework execution.

Disadvantages

- The black box nature and uncertain prediction rates.



- Training times are long, and data efficiency is low. Training data is used by neural networks to learn and increase their accuracy over time. The training process of neural networks is the focal point of determining the correct prediction of data patterns, however, it can take a long time! In fact, a successful neural network can take weeks or even months to train completely from scratch.
- Costly both financially and computationally. As previously said, neural networks are more computationally expensive than standard algorithms since they require longer to train and develop. The long duration of development also requires resources (such as human labor) to support, so using neural networks can end up being economically costly for your business.

Pre-processing for collaborative filtering

The data in the weights column used as rating was found to be skewed with the minimum weight being 1 and the maximum 352,698. To prevent bias in the result, the skewness of the weights was reduced using the percentile linearization method. This method normalizes that data according to percentiles and pushes them between 0 to 1.

The transformed weights were then converted into a rating scale from 1 to 5 using the MinMaxScaler function. MinMaxScaler subtracts the minimum value in the feature and then divides by the range. Following this process, the skewness of the weights was reduced from 43.27 to 0.0026. The weights/ratings were rounded to be integers between 1 to 5. The final data frame consisted only of the user, item(artist), and rating(weight) columns.

C. Content-based Recommendation Systems

Content-based algorithms use item features to recommend to users other items similar to what the user likes, based on their previous preferences (ratings). A matrix of predefined characteristics of items is used to recommend additional items with similar characteristics.

This project uses the TFIDF approach, and the similarity based approach both to run different models.

Under the **similarity based** approach, first the items are represented in an n-dimensional vector space and then the similarity between item vectors is computed based on cosine similarity. The matrix is then filtered for predefined nearest neighbors. The item characteristic matrix is used to fit the content based model with n nearest neighbors. The user-item-rating set is split into train and test sets and the model is fit on the train set and tested on the test set. This way predictions are made for every user-item combination⁴.

For the **TFIDF** approach, the main difference is that the tags are encoded as a weighted term vector. The term frequency measures, how often a term appears in the tags (density). This assumes that important terms appear more often e.g. metal appears in death metal, hard metal, and black metal. Since the tags are made by users and there is no specific format and can be of varying length, normalization is used. Inverse document frequency aims to reduce the weight of terms that appear in all the tags.

The **advantages** of using a content based model are that its is adaptive, and comparison between items is possible. The **disadvantages** of using content based models are that the characteristic information for each product is hard to gather and that the model can overspecialize.

Pre-processing and execution

- The tags and user_tags files were merged to create a matrix.
- For TFIDF approach, the NLtk and the Sklearn libraries were used.
- The tags were tokenized, stop words were removed, and words were stemmed.
- The TFIDF vectorizer was used to create a numerical based vector for each tag.
- All items are joined together to create a document term matrix in sparse format.
- This matrix (called df_dtm) has one row for each artist and the columns are the tags and their weights.
- The content based algorithm was then fit using 10 nearest neighbours. The model is then trained and tested on the train and test sets.
- For the second and third approach, dummy variables were made using the time data. A date column was created by combining the day, month, year columns.
- Irrelevant columns were dropped

Experiments

Since there were many tags with similar names (e.g. metal, death metal, and black metal), an attempt was made to collapse similar categories into one by using topic modeling.

We also tried topic modeling and use the topics as features, but due to systematic tags and local system memory limitations we could not implement the same. But we have the initial code in another notebook that we are submitting along with this.

The result was however not good enough to apply as some categories

⁴ Hina Hussain; Individual Assignment for Recommendation Tools

- The data was grouped by item (artistID) and the unstack function was used to pivot the tag values. This resulted in a matrix with artist ID in the rows and tag values in the columns.
- Tags which were not linked to any artist were automatically dropped.
- Special attention was paid to have one row per unique artist and that one artist can be linked to multiple tags.
- A content based model was run based on how old the tags were and another was run based on dummy variables for all tags and for the age of the tags.

D. Hybrid Recommendation Systems

Hybrid recommendation systems combine two or more recommendation strategies in different ways to benefit from their complementary advantages. The hybridization can be based on algorithms or data sources. These systems help to mitigate the issues related to individual recommendation systems and often improve predictive performance.

In this project the following hybrid recommendation systems are used:

1. Content based and Item based

- Predictions made using content based and KNN item based algorithm were extracted
- The hybrid prediction was calculated as the mean of the content based and item based prediction

2. Linear Regression using content and item based predictions

- The real ratings are used as the target variable and content based predictions and item based predictions are used as independent variables.
- The train set and test set is made using the logic in the point above
- A linear regression model is fit on the train set and predictions are made on the test set

3. Random Forest Regression using content and item based predictions

- The real ratings are used as the target variable and content based predictions and item based predictions are used as independent variables.
- The train set and test set is made using the logic in the point above
- The random forest regressor is fit using the train set. Max depth of 15 is selected and number of estimators is set to 120. A random forest is a meta estimator that fits a number of classifying decision trees on various sub-samples of the dataset and uses averaging to improve the predictive accuracy and control over-fitting. The sub-sample size is controlled with the max_samples parameter if bootstrap=True (default), otherwise the whole dataset is used to build each tree.
- The test set was used to make the predictions.

4. Adaboost using content and item based predictions

- The real ratings are used as the target variable and content based predictions and item based predictions are used as independent variables.
- The train set and test set is made using the logic in the point above
- The adaboost regressor is used to train the model. Under this approach trees are grown sequentially using information from previously grown trees. Adaboost regressor is a meta-estimator that begins by fitting a regressor on the original dataset and then fits additional copies of the regressor on the same dataset but where the weights of instances are adjusted according to the error of the current prediction. As such, subsequent regressors focus more on difficult cases.
- In this case, the random forest regressor was used as the base estimator and the number of estimators was set to 500

5. Bayesian Ridge using content and item based predictions

- The real ratings are used as the target variable and content based predictions and item based predictions are used as independent variables.
- The train set and test set is made using the logic in the point above
- The main difference between Bayesian ridge and ordinary least square is that the coefficient weights are slightly shifted toward zeros, which stabilises them. The estimation of the model is done by iteratively maximizing the marginal log-likelihood of the observations.
- In this case the Bayesianridge model is used with shape parameter for the Gamma distribution set to 0.0012. The normalize option is set to true which means the regressors X will be normalized before regression by subtracting the mean and dividing by the l2-norm.

E. Evaluation of all Models

All the models applied in this project were evaluated using different accuracy measures.

| Algorithm | RMSE | MAE | Recall | Precision | F1 |
|---------------------------------------|--------|--------|--------|-----------|--------|
| Hybrid with AdaBoost | 72.10% | 51.60% | 49.90% | 98.00% | 66.10% |
| SVD ++Tuned | 77.60% | 62.20% | 33.40% | 94.50% | 49.30% |
| SVD | 74.20% | 58.00% | 39.10% | 94.20% | 55.30% |
| KNNBaseline | 79.70% | 63.70% | 34.00% | 91.60% | 49.60% |
| BaselineOnly | 78.10% | 63.70% | 22.50% | 97.20% | 36.50% |
| Hybrid with Random Forest | 72.10% | 53.70% | 41.40% | 98.50% | 58.30% |
| SVD++ | 72.90% | 56.40% | 42.60% | 94.00% | 58.60% |
| Content Based (Using recency) | 75.20% | 57.70% | 44.80% | 93.00% | 60.40% |
| Hybrid using Gaussian NB | 82.70% | 63.40% | 0.00% | 0.00% | 0.00% |
| Hybrid - Linear Regression | 82.70% | 63.40% | 38.70% | 93.90% | 54.80% |
| KNNWithMeans | 85.60% | 65.50% | 41.10% | 90.40% | 56.50% |
| Content_based (with recency and tags) | 91.50% | 70.50% | 34.80% | 93.10% | 50.60% |
| Content_based (Tags only) | 91.60% | 70.50% | 34.70% | 93.20% | 50.60% |

The accuracy measures used are described below⁵:

1. Root Mean Squared Error (RMSE)

RMSE is a prediction accuracy method. It measures the standard deviation of the prediction error. The prediction error is the difference between the prediction and the true result. RMSE is calculated by taking the mean of all the square root of the prediction errors. Therefore, a lower RMSE indicates better predictions.

2. Mean Absolute Error (MAE)

MAE is also a prediction accuracy method. MAE of a model is the mean of the absolute values of the individual prediction errors over all instances in the test set. The prediction error is the difference between the prediction and the true result. A low MAE indicates better predictions.

3. Normalized Discounted cumulative gain (NDCG)

NDCG is a ranking accuracy metric. Normalized discounted cumulative gain is discounted cumulative gain divided by idealized discounted cumulative gain. This measure gives the result between 0 to 1, where 1 is relevant and 0 is not relevant. For discounted cumulative gain, this is calculated by the accumulated of all the logarithmically reduced real ranking position, whereas idealized discounted cumulative gain is calculated by the accumulated of all the ranking position if they give the highest value after the logarithmic reduction factor.

4. F1-Score (classification accuracy)

F1 score is a classification accuracy metric. It is the harmonic mean of the precision and recall (representing exactness and completeness) which are used in binary classification. Precision represents how many items selected are relevant; it is computed by true positives divided by the sum of true positives and false positives. Recall represents how many relevant items are correctly identified as true positives; this can be computed by true positives divided by the sum of the sum of true positives and false negatives. Higher F1-Score shows higher accuracy of the model.

F. Selected Model and its Hyperparameter tuning using Grid Search

Judging from the various evaluation methods discussed in the section above, the SVD++ model was chosen as the final model as it produced the best performance. Grid search was used to find the optimal hyperparameters of the model which result in the most 'accurate' predictions. A grid search with a range of options for the hyperparameters was run for the SVD++ model. The search was run on 4 hyperparameters – number of factors, number of epochs, learning rate, and regularization term.

⁵ Mui Han Ma; Individual Assignment for Recommendation Tools

G. Qualitative assessment of the results from the final model

Our final recommendation system generates 10 recommended artists based on: (1) 5 recommendations based on AdaBoost Hybrid System and (2) 5 recommendations based on SVD++ CF System.

In order to illustrate the whole recommendation process, User #504 was selected to walk through the process.

The following table shows the artist tag of the artists which listened by User #504:

According to the artist tag, this user usually listened to pop, dance, acoustic, core and electronic music.

| userid | artistid | name | artist_tag |
|--------|----------|------------------------|--|
| 0 | 504 | 289 Britney Spears | nostalgia,dance,pop,pop,bitch,dance,90s,its,britney,bitch,legend,fave,songs,ever,fuck,yeah,pop,dance,soundtrack,of,my,life,90s,electropop,why,on,earth,want,this,release,d,all,time,hi,but,the,party,dont,stop,for,haters,celebrities,who,one,day,im,determined,to,meet,new,girls,did,one,of,my,favorites,was,among,female,vocalists,legend,female,vocalist,90s,love,at,first,listen,awesome,kali,gym,justin,timberlake,party,line,among,i,was,fighting,with,myself,not,to,identify,with,but,i,love,bitch,ring,britney,spears,2000,female,vocalists,we,should,be,together,pop,underrated,dance,let's,dance,love,at,first,listen,fun,addictive,cover,90s,electronic,so,hot,stand,out,beyond,pop,female,dance,pop,female,female,dance,pop,dance,female,vocalists,electronic,(3,catchy,90s,dance,for,days,when,i,am,particularly,angry,with,you,dance,dance,music,90's,songs,i,can,actually,listen,to,on,repeat,chrismas,why,on,earth,is,this,just,a,bonus,track,love,at,first,listen,gross,meaning,but,good,song,pop,female,vocalist,love,at,first,listen,electronic,pop,very,sexy,pop,female,vocalists,dance,pop,pop,dance,female,vocalists,legend,britney,spears,circus,dance,pop,female,vocalists,dance,pop,female,vocalists,among,legend,90s,female,pop,90s,jun,usa,love,sex,pop,dance,female,vocalists,(best,bitch,songs,my,favorites,songs,diva,diva,why,on,earth,is,this,just,a,bonus,track,dance,sufco,cute,awesomesauce,female,vocalists,pop,princess,d,op,legendaria,linda,linda,female,vocalists,pop,dance,female,vocalist,female,vocalists,singer-songwriter,fun,female,sexy,gully,pleasure,female,vocalists,legend,pop,legends,godney,britney,spears,bald,party,catchy,gully,pleasures,pop,dance,female,vocalists,addictive,catchy,as,hell,favorite,song,(3,female,vocalists,dance,(3,pop,queen,supersal,grammy,winner,recordman,riv,unreal,leaves,living,legend,blackout,truth,transcendent,pop,princess,d,iva,britney,spears,favorite,pop,environment,britney,spears,bitch,90s,britney,spears,dance,britney,pop,legend,dont,hot,queen,of,pop,sex,combi,did,performer,best,seller,pop,icon,seventeen,among,first,best,song,catchy,100,female,vocalists,dance,britney,spears,legend,female,vocalists,dance,pop,favorites,legend,acoustic,rnb,electronic,pop,female,vocalist,balad,90s,gully,pleasures,american,tad,999,love,song,pop,female,vocalists,rock,dance,alternative,hq,pop,female,rnb,pop,princess,love,song,love,britney,spears,favorite,britney,spears,the,best,ever,blonde,cousin,legend,pop,electronic,as,an,tags,acappella,pop,huge,artists,princess,of,pop,legend,2000,play,back,90s,pop,rock,female,vocalist,dance,pop,playback,best,songs,of,90s,best,chrismas,songs,xmas,girls,top,20,songs,dana,sang,on,a,factor,90s,female,vocalists,pop,electronic,angels,circus,pop,female,vocalists,beautiful,sexy,pop,the,best,race,american,legend,soundtrack,female,female,vocalists,pop,like,crack,for,my,soul,sugato,homosexual,handclaps,dance,catchy,american,dance,sexy,balad,pop,pop,sexy,pop,dance,hot,dance,female,vocalist,american,music,female,vocalists,britney,rulzz,rnb,legend,sexy,urban,pete,dink,and,my,female,falls,britney,rulzz,awesome,90s,pop,electronic,pop,pop,pop,female,vocalists,catchy,100,female,vocalists,legend,female,vocalists,dance,pop,electronic,fav,tracks,for,female,singers,catchy,diva,favorites,albums,i,own,i,love,favorites,awesome,sexy,among,electronic,electronic,gully,pleasure,love,songs,the,did,i,again,balad,catchy,beautiful,(3,female,vocalist,dance,pop,addictive,britney,spears,pop,britney,spears,female,female,vocalists,dance,pop,dance,pop,dance,legend,pop,circus,amd,love,at,first,listen,queen,of,pop,britney,spears,girl,power,dance,legend,pop,britney,pop,princess,real,que,fal,later,hq,playback,pop,britney,spears,dance,hot,legend,among,perfect,bitch,best,remix,ever,its,britney,bitch,pop,rock,dance,love,among,female,vocalists,female,vocalists,pop,britney,spears,bitch,hq,real,playback,awesome,inside,inside,britney,spears,sexy,pop,dance,female,vocalists,american,sexy,best,too,40,beyond,godney,relations,to,get,up,in,the,morning,my,my,impulse,happy,pop,female,vocalists,living,legend,electropop,pop,10s,melodic,drug,god,dance,female,vocalist,electronic,the,best,queen,legend,female,vocalists,break,the,ice,pop,britney,spears,dance,soundtrack,d,iva,queen,of,pop,legend,female,vocalist,rnb,pop,living,legend,anah,pop,perfect,rnb,britney,the,singles,collection,90s,awesome,pop,pop,britney,spears,pop,2003,2007,2009,catchy,balad,among,sexy,female,2004,female,vocalists,female,vocalists,queen,of,pop,hot,beautiful,american,90s,pop,legend,dance,legend,pop,legend,britney,spears,dance,perfect,masterpiece,2003,female,female,sexy,fun,favorites,remember,the,time,for,britney,spears,legend,female,vocalists,remix,pop,music,pop,pop,electronic,dance,female,vocalists,dance,female,vocalists,dance,legend,pop,pop,pop,sexy,love,30,pop,brandy,janel,polson,pop,vocalists,feminine,favorite,did,but,god,legend,american,legend,awesome,great,good,best,ever,pop,infectious,max,martin,britney,spears,pop,blackout,legend,britney,spears,pop,britney,spears,queen,dmxy,amazing,voice,legend,love,karmilla,pop,female,vocalists,britney,spears,circus,living,legend,perfect,godney,masterpiece,hot,queen,of,pop,female,pop,princess,pop,princess,circus,britney,womanizer,great,dances,girls,godney,luxo,(3,legend,american,pop,covers,among,sexy,pop,singles,sexy,beautiful,legend,pop,pop,favorites,amazing,pop,rock,female,vocalists,rnb,american,electropop,catchy,hot,dance,(3,d,iva,legend,90s,beautiful,90s,dance,dancepop,love,party,usa,dance,dance,pop,bem,pop,awesome,sexy,among,american,icon,90s,queen,of,pop,pop,rock,amazing,awesome,favorite,artists,90s,weekly,top,artists,sexy,sexy,lively,90s,did,best,spears,electronic,pop,love,bitch,d,dance,pop,electronic,britney,spears,remix,super,electronic,female,makes,me,wanna,dance,all,time,favorites,weekly,top,tracks,better,than,lady,gaga,but,i,love,too,dearly,crimes,amotional,summer,song,urban,sensual,legend,legend,comeback,the,best,thing,that,ever,happened,to,music,legendary,circus,living,legend,pop,female,britney,spears,pop,rock,country,lain,90s,cool,female,90s,jun,pop,dance,female,vocalists,legend,awesome,addition,best,singer,ever,best,singer,ever,2003,over,amazing,love,at,first,listen,dance,female,vocalists,bitch,amazing,2000,dance,pop,pop,(3,female,vocalist,dance,pop,legend,pop,god,godney,legendary,low,dmxy,2003,dmxy,pop,circus,britney,spears,pop,pop,female,vocalists,pop,legend,britney,spears,pop,od,de,britney,en,el,auto,con,repas,para,carfala,con,for,recuerdos,de,la,infancia,con,asis,Hiero,flo,crush,american,bitch,jany,d,op,amou,meu,d,iva,legend,love,if,masterpiece,glam,sexy,songs,best,of,2007,punk,rock,female,vocalists,britney,spears,love,legend,did,britney,spears,teen,pop,electro,pop,dance,pop,britney,spears,bitch,spears,girl,90's,seen,live,cute,unreleased,2007,songs,to,male,flatchuzch,zr,songs,that,remind,me,of,summer,07-09,th,so,catchy,blackout,britney,spears,pop,dance,covers,queen,of,pop,teen,pop,gay,icon,gay,gay,icon,female,disco,hot,female,vocalists,seen,live,alternative,rnb,dance,sexy,girl,fun,pop,amou,best,of,pop,covers,pop,favorites,princess,electronic,legend,pop,my,favorites,music,favorites,bitch,top,wilds,jun,dance,dive,dance,bitch,rnb,bitch,catchy,female,vocalists,britney,spears,legend,singer-songwriter,female,vocalists,britney,spears,dance,pop,dance,female,vocalists,sexy,gossip,girl,cannibal,pop,sexy,pop,who,should,be,sad,great,memories,with,my,sis,britney,perfect,love,melodies,britney,spears,catchy,sexy,female,vocalists,dance,dmxy,queen,louella,popular,dance,35w,legend,female,vocalists,female,vocalists,love,american,d,iva,the,best,electro,legend,rnb,dance,pop,electronic,women,female,vices,britney,spears,britney,spears,pop,dance,rnb,dmxy,girl,sexy,pop,dance,pop,love,let,me,be,bitch,rich,and,famous,pop,rock,me,in,pieces,of,bitch,bitch,bitch,90s,lady,pop,dance,favorites,d,iva,sexy,female,vocalist,pop,awesome,songs,songs,belief,among,90s,pure,my,britney,dance,need,i,was,remind,pop,lively,britney,spears,dance,britney,spears,pop,gully,pleasures,female,vocalists,pop,dance,dance,american,lean,then,d,sexy,electronic,pop,dance,dance,sexy,pop,female,female,vocalist,female,vocalist,legend,d,stop,vocalists,dance,pop,dance,music,prim,awesome,electronic,pop,pop,balad,90s,90s,pop,dance,legend,90s,icon,catchy,electropop,luxo,j,crisis,pop,pop,britney,spears,90s,living,legend,american,male,version,hot,electronic,britney,female,vocalist,sexy,pop,perfect,queen,queen,legend,queen,of,pop,the,best,pop,perfect,britney,spears,electropop,electronic,addition,seen,live,womanizer,american,homotic,electronic,d,iva,remix,pop,bitches,cover,electronic |
| 1 | 504 | 1676 BrokeNCYD E | but,the,party,dont,stop,these,gays,fucking,rule,awesomesauce,apic,cancer,than,this,billing,music,crime,against,humanity,(3,scream,crusike,sh,hq,hop,scream,electronic,calleente,scream,hq,bounce,electronic,gabage,you,ears,will,bleed,bc,d,motherfucker,cure,rock,forreal,scream |
| 2 | 504 | 1680 The Pheary Set | powerpop,fun,chrismas,pure,w,electronic,awesomesauce,cute,catchy,male,vocalists,dance,apic,powerpop,male,vocalists,pop,pop,rock,cute,(3,techno,pop,electronic |






(1) Recommendations based on AdaBoost Hybrid System

The first 5 recommendations are based on the AdaBoost Hybrid System which are having the highest similarity to the tags from the User #504. The recommended artists also have tags related to pop, dance and electronic music as highlighted.











| | image_url | name | url | artist_tag |
|---|-----------|---------------------------|---|---|
| 0 | | Cordell do Fogo Encantado | http://www.last.fm/music/Cordell+do+Fogo+Encantado | multiculti,brazilian,brasil,guitar,mpb,electic,alternative,indie,brazilian,traditional,multicultural,multikulti,seen live,alternative,macumba,music,percussion |
| 1 | | Lucky Twice | http://www.last.fm/music/Lucky+Twice | swedish,female,vocalists,dance,swedish,dance,electronica,pop,electronic |
| 2 | | Air | http://www.last.fm/music/Air | 90s,chillout,electronic,french,downtempo,electronic,french,ambient,trip-hop,french,80s,chillout,alsolike,friends,ambient,favorite artist,electronic,electronic,trip-hop,alternative,sexy,dream,pop,downtempo,ambient,electronic,france,electronic,electronic,electronica,chillout,electronic,french,electronica,alternative,indie,electronic,ambient,electronica,chillout,chillout,indietronica,90s,indie,pop,instrumental,french,jazz,dance,electronic,electronica,electronic,trip-hop,chillout,ambient,downtempo,alternative,electronica,dub,mellow,electronic,electronic,mellow,instrumental,sexy,english,chillout,ambient,electronic,electronica,seen live,air,electronic,ambient,downtempo,trip-hop,vocal,mellow,electronic,alternative,electronic,ambient,sexy,dreamlike,chillout,favourite,songs,love,chillout,chillout,ambient,electronic,french,favorites,brilliant,trip-hop,curious,about,classic,rock,pastel,my electro feelings,jove,chillout,donbeat,efly,demo,ambient,electronic,electronic,soundtrack,chillout,chillout,ambient,electronic,trip-hop,ambient,electronic,electronic,dream,indie,ambient,love it,trip-hop,atmospheric,electro,veronica,mas,lounge,downtempo,ambient,downtempo,chillout,meditative,hypnotic,dark,instrumental,experimental,trip-hop,electronic,european,aana,electronic,trip-hop,electronic,sweet,electronic,gay,ambient,electronic,chillout,electronic,lounge,alternative,chillout,cool,alternative,music,for my soul,ambient,electronic,90s,sad,brilliant,electronic,amazing,relax,beautiful,music,sweet,autumn,beautiful,sex,music,relaxing,indie,piano,dreamy,french,chillout,electronic,electronica,downtempo |
| 3 | | Gary Numan | http://www.last.fm/music/Gary+Numan | cyberpunk,new wave,electronic,industrial,synthpop,80s,electronic,new wave,ultimate favorite,new wave,80s,awesome,electronic,80s,gay,uman,new wave,best,electronic,80s,favorite artists,electronic,my favorites,electronic,alternative,random,songs,1983 songs,male vocalists,1985 songs,80s,darkwave,80s,talking heads,1984 songs,new romantic,synth,alternative,dance,fav songs,90's,synthpop,synthpop artists,deep-clash,1980 songs,dazs favourites,80s,synthpop,jam hot,songs to listen to on a sunny day,80s,new wave,new wave songs,fav artists,industrial,new wave,synthpop,new wave,creepy,favorite tracks,new wave,synthpop songs,1982 songs,british artists,1981 songs,new wave artists,darkwave |
| 4 | | Cause & Effect | http://www.last.fm/music/Cause%2B%2526%2BEffec | pop,i love 90s,80's,synth pop,synth,electronic,dance,synth wave,new wave,synthpop,synthpop,electronica,classic synthpop,synthpop,90s,electronic,synth pop,awesome |

(2) Recommendations based on SVD++ CF System for variety

Another 5 recommendations are based on the best CF model, which is the SVD++ model for variety. In order to recommend the 5 recommendations, the model selects the top 5 artists in which the user listens to similar artists. Same as Hybrid System, these recommended artists also have tags related to pop, dance, acoustic, core and electronic music.

| image_url | name | url | artist_tag |
|---|------------------|---|--|
|  | Trap Them | http://www.last.fm/music/Trap+Them | noisecore ,mathcore ,hardcore ,crust ,chaotic hardcore ,hardcore ,00s ,hardcore ,hardcore ,grindcore ,grindcore ,mathcore |
|  | The Books | http://www.last.fm/music/The+Books | nostalgia ,check ,downtempo group favourites ,experimental ,indietronica ,electronic ,experimental ,electronica ,awesome-ness ,instrumental ,experimental |
|  | Mat Kearney | http://www.last.fm/music/Mat+Kearney | acoustic ,male vocalists ,acoustic ,alternative ,acoustic ,acoustic ,singer-songwriter ,life unexpected ,ambient ,indie ,singer-songwriter ,alternative ,one tree hill |
|  | Massari | http://www.last.fm/music/Massari | rnb ,4m4zinq ,f4nt4stic ,pop rock ,specials to 3mmy ,dance |
|  | Declan Galbraith | http://www.last.fm/music/Declan+Galbraith | touching ,inspiring |

At the end, all these 10 recommendations are combined to a list to the User #504 as the final recommendations.

| image_url | name | artist_tag |
|---|----------------------------------|---|
|  | Cordell do Fogo Encantado | multiculti ,brazilian ,brasil ,guitar ,mpb ,eclectic ,alternative ,indie ,brazilian ,traditional ,multicultural ,multikulti ,seen live ,alternative ,macumba music ,percussion |
|  | Lucky Twice | swedish ,female vocalists ,dance ,swedish ,dance ,electronica ,pop ,electronic |
|  | Air | 90s ,chillout ,electronic ,french ,downtempo ,electronic ,french ,ambient ,trip-hop ,french ,80s ,chillout ,alsolike ,friends ,ambient ,favorite artist ,electronic ,electronic ,trip-hop ,alternative ,sexy ,dream pop ,downtempo ,ambient ,electronic ,electronic ,france ,electronic ,electronic ,electronica ,chillout ,electronic ,french ,electronica ,alternative ,indie ,electronic ,ambient ,electronica ,chillout ,chillout ,indietronica ,90s ,indie pop ,instrumental ,french ,jazz ,dance ,electronic ,electronica ,electronic ,trip-hop ,chillout ,ambient ,downtempo ,alternative ,electronica ,dub ,mellow ,electronic ,electronic ,mellow ,instrumental ,sexy ,english ,chillout ,ambient ,electronic ,electronica ,seen live ,air ,electronic ,ambient ,downtempo ,trip-hop ,vocal ,mellow ,electronic ,alternative ,electronic ,ambient ,sexy ,dreamlike ,chillout ,favourite songs ,love ,chillout ,chillout ,ambient ,electronic ,french ,favourites ,brilliant ,trip-hop ,curious about ,classic rock ,pastel ,my electro feelings just chillin ,downbeat ,effy ,demo ,ambient ,electronic ,electronic ,soundtrack ,chillout ,chillout ,ambient ,electronic ,trip-hop ,ambient ,electronic ,electronic ,dream ,indie ,ambient ,love it ,trip-hop ,atmospheric ,electro ,veronica mars ,lounge ,downtempo ,ambient ,downtempo ,chillout ,meditative ,hypnotic ,dark ,instrumental ,experimental ,trip-hop ,electronic ,european ,aana ,electronic ,trip-hop ,electronic ,sweet electronic ,gay ,ambient ,electronic ,chillout ,electronic ,lounge ,alternative ,chillout ,cool ,alternative ,music for my soul ,ambient ,electronic ,90s ,sad ,brilliant ,electronic ,amazing ,relax ,beautiful music ,sweet ,autumn ,beautiful ,sex music ,relaxing ,indie ,piano ,dreamy ,french ,chillout ,electronic ,electronica ,downtempo |
|  | Gary Numan | cyberpunk ,new wave ,electronic ,industrial ,synthpop ,80s ,electronic ,new wave ,ultimate favorite ,new wave ,80s ,awesome ,electronic ,80s ,gary numan ,new wave ,best ,electronic ,80s ,favorite artists ,electronic ,my favorites ,electronic ,alternative ,random songs ,1983 songs ,male vocalists ,1985 songs ,80s ,darkwave ,80s ,talking heads ,1984 songs ,new romantic ,synth ,alternative dance ,fav songs ,90's ,synthpop ,synthpop artists ,deep-clash ,1980 songs ,dazs favourites ,80s ,synthpop ,jam hot ,songs to listen to on a sunny day ,80s ,new wave ,new wave songs ,fav artists ,industrial ,new wave ,synthpop ,new wave ,creepy ,favorite tracks ,new wave ,synthpop songs ,1982 songs ,british artists ,1981 songs ,new wave artists ,darkwave |
|  | Cause & Effect | pop ,i love 90s ,80's ,synth pop ,synth ,electronic ,dance ,synth wave ,new wave ,synthpop ,synthpop ,electronica ,classic synthpop ,synthpop ,90s ,electronic ,synth pop ,awesome |
|  | The Books | nostalgia ,check ,downtempo group favourites ,experimental ,indietronica ,electronic ,experimental ,electronica ,awesome-ness ,instrumental ,experimental |
|  | Vision System | swedish |
|  | Brutality Will Prevail | NaN |
|  | Declan Galbraith | touching ,inspiring |
|  | Frankie Valli & The Four Seasons | rock ,catchy ,60s ,popular ,pop |

References:

1. K Nearest Neighbors
<https://www.fromthegenesis.com/pros-and-cons-of-k-nearest-neighbors/>
2. SVD++
<http://www.diva-portal.org/smash/get/diva2:1439511/FULLTEXT01.pdf>
3. CoClustering
<https://en.wikipedia.org/wiki/Biclustering>
4. RMSE
https://en.wikipedia.org/wiki/Root-mean-square_deviation
5. MAE
https://en.wikipedia.org/wiki/Mean_absolute_error
6. NDCG
https://en.wikipedia.org/wiki/Discounted_cumulative_gain
7. F1-Score
<https://en.wikipedia.org/wiki/F-score>
8. Random Forest

<https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestRegressor.html>
9. Adaboost

<https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.AdaBoostRegressor.html>
10. Bayesian Ridge

https://scikit-learn.org/stable/auto_examples/linear_model/plot_bayesian_ridge.html
11. Neural Networks
<https://www.ibm.com/cloud/learn/neural-networks>