

A Probabilistic Topic Models Based Music Recommendation System

Lixuan Zhu

Rutgers University-New Brunswick

lz306@scarletmail.rutgers.edu

Yi Zhong

yz614@scarletmail.rutgers.edu

Zheng Yu

zy120@scarletmail.rutgers.edu

Shihao Su

ss2719@scarletmail.rutgers.edu

Abstract

We will implement a music recommendation system utilizing Probabilistic Topic Models. Specifically, given a song, the system will find a series of similar songs according to features such as metadata, tags and acoustic characteristics.

1. Introduction

Topic Models are a collection of algorithms that finds the underlying topics from a large amount of text documents. According to Blei, each document is a mixture of corpus-wide topics; each topic is a distribution over words and each word is drawn from one of those topics [1]. When a distribution of topics of the document emerge from topic modeling, we can compute the similarity between two documents by comparing their topic distributions, thus the documents with high similarity to the given document can be recommended to user.

There are two general approaches when designing recommendation systems: collaborative filtering and content-based filtering, each with their strengths and weaknesses. Collaborative filtering utilizes the preferences of existing users and predict whether a specific user will like an item based on the decision of similar users. Music streaming services such as Last.fm utilizes collaborative filtering to recommend songs based on specific user profiles [6]. However, a natural challenge for collaborative filtering methods is the lack of user ratings, also referred to as the cold start problem [6]. For example, when a new song is published, there is not enough user rating for the system to know which users will like the song. The other method, content-based filtering, focuses on the similarity of the songs itself. Pandora, another popular music streaming service, utilized the metadata(artist, genre, etc.) and tags to find similar songs

given an initial seed song [4]. Although content-based approaches require very little initial information, it is limited to recommend the songs similar to the original seed since the information about the song is limited.

In this project, we take the content-based filtering a step forward. We utilize two probabilistic topic models, latent dirichlet allocation(LDA) and hierarchical dirichlet process(HDP), to discover the similarity between songs and compare their performances. We consider each song as a document and the features of the song as words. It is important to define what we mean by features. Traditional recommendation systems uses the metadata of the song itself and tags by users as measures for similarity. In our model, we incorporate acoustic properties of the music such as pitch, tempo, verses and chords used, mood progression, etc. The timeline API from Gracenote is used to retrieve acoustic properties from the songs[5]. Collectively, we consider the traditional measures and acoustic properties of the song as features. Thus each feature is a word in the document. The probabilistic topic models process the entire collection of songs as a corpus. The advantage of utilizing acoustic properties is the range of songs recommended can be extended. Songs with similar acoustic properties may not be in the same genre or tagged with similar attributes. The details of the implementation of our models and the challenges of utilizing acoustic properties will be covered in section 3.

Another focus of our project is to compare two topic modeling algorithms, LDA and HDP. We plan to analyze the performance of both algorithms in terms of complexity and recommendation result. In comparison, our HDP model is based on LDA and more complex in terms of implementation. However, LDA requires a priori entry of the number of topics in the corpus. How to choose the input value and whether the recommendations are meaningful will be addressed in section 4.

2. Prior Work

Blei, Ng, and Jordan specified a popular probabilistic model for discovering latent semantic topics in large collections of text documents in Latent Dirichlet Allocation[2]. Hu and Saul used symbolic and audio model of music to classify western tonal music in A Probabilistic Topic Model for Music Analysis[3]. Hoffman, Blei and Cook developed a HDP-based way to compute the similarity of songs, which is much faster and better than previous approaches that compare single Gaussian distributions directly[7]. Hariri, Mobasher and Burke presented a context-aware music recommender system which infers contextual information based on the most recent sequence of songs liked by the user[8]. Salit, Weinshall and Chechik applied the topic model to model music influence and study properties of influence[9].

3. Models, Assumptions and Requirements

3.1. Word Model of Songs

In order to find topics from songs with probabilistic topic model, we need to preprocess the audio and metadata of the songs to extract audio words. Based on the probabilistic topic model of Diane J. Hu and Lawrence K. Saul[3], our model is defined as follow:

1. Word list is defined as $u \in \{\{\text{metadata}\}, \{\text{tag}\}, \{\text{pitch}\}, \{\text{tempo}\}, \{\text{verses}\}, \{\text{chords}\}, \{\text{mood progression}\}\}$. Every subset in u (e.g. metadata, tag, etc) will be refined to detailed audio words during in the process. Words in different subsets will be weighted as we don't want artist, genre type of words to have too much contribution to the final results;

2. A Segment is defined as a group of words in an unit of time. The words in the n th segment is defined by $\text{sgn} = \text{sgn}_1, \dots, \text{sgn}_L$, where sgn_l is the l th word in the segment.

3. A song s is defined as a sequence of D segments: $s = \text{sg}_1, \dots, \text{sg}_D$.

4. The collection of M songs are $S = s_1, \dots, s_M$.

5. A topic $\beta \in \{1, \dots, K\}$ is a probability distribution over the elements of word list.

3.2. Latent Dirichlet Allocation

We know that LDA is an unsupervised, probabilistic, generative model for discovering latent semantic topics in large collections of document data. Now we adapt the LDA for songs. Comparing with original LDA, our model treats each song as a document and the features of a song as words. Our purpose of modeling is to mining the topics in collections of song dataset. First we have following assumption: Each discovered topic is characterized by its own distribution over features. Each song is then characterized as a random mixture of topics indicating the proportion of

features each topic over songs[2]. Our goal is exactly such random mixture of topics, which expresses the meaning and style of a song. By using this, we can compare how similar one song is to another by calculating how similar the corresponding topic mixtures are and recommend similar songs. When we do generative process of LDA, we need specify topics we will use first, then calculate the probability of the feature generate by the topic. Finally we use Gibbs sampling to go over all the features. In order to do this, we describe LDA more formally with following notation. Our model suppose each song corresponding with K topics $\beta_{1:k}$. The topic proportions for the d th song are θ_d , the topic assignment for d th song are Z_d , the observed features for song d are W_d . By using these notation, the joint distribution of the generative process for LDA is shown as equation 1. Then we compute the conditional (posterior) distribution of the topic, which shown as equation 2, and the denominator is the marginal probability of observed corpus. Finally the two distribution will converge and we get what we want.

$$p(\beta_{1:k}, \theta_{1:D}, z_{1:D}, w_{1:D}) \\ = \prod_{i=1}^K p(\beta_i) \prod_{d=1}^D p(\theta_d) \\ \left(\prod_{n=1}^N p(z_{d,n} | \theta_d) p(w_{d,n} | \beta_{1:k}, z_{d,n}) \right). \quad (1)$$

$$p(\beta_{1:k}, \theta_{1:D}, z_{1:D} | w_{1:D}) \\ = \frac{p(\beta_{1:k}, \theta_{1:D}, z_{1:D}, w_{1:D})}{p(w_{1:D})}. \quad (2)$$

3.3. Hierarchical Dirichlet Process

The Hierarchical Dirichlet Process (HDP) is a model of grouped data. We will construct a HDP to model songs represented as a collection of features. The HDP will generate the number of clusters; the clusters here represent the different topics of songs. Rather than associate each song with a single topic in local, each song is represented as a group of global topics. Thus, each song is represented as a distribution over latent components, but the population of latent components is shared across songs. The HDP model of songs can be generated with the Chinese Restaurant Franchise (CRF). The CRF takes two hyper parameters α and γ . Each song j has its own CRP, and each feature vector chooses a topic from $\text{CRP}(\alpha)$. If it fit a new distribution, then it chooses a topic for that distribution from a global CRP (with hyper parameter γ) shared by all songs that is, it either chooses a topic that is already being assigned to some other distributions m with probability proportional to m , or it chooses a new topic with probability proportional to γ . In this setup, the songs correspond to groups and topics

correspond to the factor θ_{ji} . We also let $\phi_1, \phi_2, \dots, \phi_k$ denote k random variables distributed according to H ; this is the global list of topics. We also introduce variables, ψ_{jt} , that represent the distribute-specific choice of topics; in particular, ψ_{jt} is the topic assigned to distribution t in song j [10]. We put these factors into HDP model and get the formula below.

$$\theta_{ji} | \theta_{j1}, \dots, \theta_{j,i-1}, \alpha_0, G_0 \sim \sum_{i=1}^{m_j} \frac{n_{ji}}{i-1+\alpha_0} \delta_{\psi_{ji}} + \frac{\alpha_0}{i-1+\alpha_0} G_0.$$

Finally, we cluster the data with Collapsed Gibbs sampling algorithm in the HDP model. The clusters we get are the specific global topics of all songs. The number of clusters K can be the input of LDA model.

4. Evaluation

The goal of our system is to recommend songs for users. Therefore, the most important evaluation will be the accuracy of our recommendation. In other word, are the songs selected by our system satisfy the emotional demands of users? We will evaluate the recommendation result based on feedback from users.

Moreover, we will implement two algorithms (LDA and HDP) and test them on specific dataset. We used data set from Gracenote, Inc[5] to extract acoustic features of the songs. To evaluate how our two models fit the data, we need to compare the topics extracted between the two models. LDA requires a priori entry of cluster numbers K . We will input different K to evaluate the output and decide the range of K for the best output.

5. Plan

The project will be finished by early December. The first step is to evaluate the existing implementation of LDA and HDP algorithms and see if any of them can be tweaked to fit our data. This step will be finished by Nov. 6th. The next step will be to explore ways to measure similarities between songs. When we have a distribution of features of the songs, how much weight do we put on each feature and how do we calculate the similarities? This step will be done by Nov. 13th. The third step will be building user interface to interactively make recommendations. This step will be done by Nov. 27th. The last step will be testing the system, evaluating the algorithms and writing project report. This step will be finished by Dec. 3th.

References

- [1] D. M. Blei. Probabilistic topic models, 2011. <https://www.cs.princeton.edu/blei/kdd-tutorial.pdf>. 1

- [2] M. I. J. David M. Blei, Andrew Y. Ng. Latent dirichlet allocation. *Journal of Machine Learning Research*, 2(1), 2003. 2
- [3] L. K. S. Diane J. Hu. A probabilistic topic model for music analysis, 2009. NIPS Workshop on Applications for Topic Models: Text and Beyond. 2
- [4] Y. Hu. *A Music Recommendation System Based on User Behaviors and Genre Classification*. PhD thesis. 1
- [5] G. Inc. Entertainment metadata, 2016. <https://developer.gracenote.com/metadata>. 1, 3
- [6] N. R. M. Elahi, F. Ricci. A survey of active learning in collaborative filtering recommender systems. *Computer Science Review*, 20(1):29–50, 2016. 1
- [7] P. C. M. Hoffman, D. Blei. A probabilistic topic model for music analysis, 2008. Proceedings of the 9th International Conference on Music Information Retrieval. 2
- [8] R. B. N. Hariri, B. Mobasher. Context-aware music recommendation based on latent topic sequential patterns, 2008. 2
- [9] G. C. U. Shalit, D. Weinshall. Modeling musical influence with topic models, 2013. 2
- [10] e. a. Y. Teh. Hierarchical dirichlet processes. *Journal of the american statistical association*. 3