Allegheny College
Department of Computer Science

Senior Thesis

Music Recommendation by Mapping Music and Descriptive Paragraph

by

Xingbang Liu

ALLEGHENY COLLEGE

COMPUTER SCIENCE

Project Supervisor: **Janyl Jumadinova**
Co-Supervisor: **Gregory M. Kapfhammer**

December 11, 2019

Abstract

Provide a concise summary of your research project of approximately 250 words. Remember that the abstract is *not* an introduction, it is a *summary* of the entire document, including the results and future direction of the project.

Acknowledgment

Theses should acknowledge assistance received in any of the following areas:

- Designing the research
- Executing the research
- Analyzing the data
- Interpreting the data/research
- Writing, proofing, or copyediting the manuscript

Abbreviations

All abbreviations used in the thesis should be listed here, with their definitions, in alphabetical order. This includes trivial and commonly used abbreviations (at your own discretion), but not words that have entered into general English usage (such as laser or DNA). In particular, non-standard abbreviations should be presented here. This is an aid to the reader who may not read all sections of the thesis.

PPT	positive partial transpose
SRPT	Schrödinger-Robertson partial transpose

Glossary

Dipole Blockade	Phenomenon in which the simultaneous excitation of two atoms is inhibited by their dipolar interaction.
Cavity Induced Transparency	Phenomenon in which a cavity containing two atoms excited with light at a frequency halfway between the atomic frequencies contains the number of photons an empty cavity would contain.

If desired, an optional and short dedication may be included here.

Contents

List of Figures

List of Tables

Template Overview

You should first modify the documents in the preamble, things that appear before the main text as detailed below.

Front page: use the one provided in this template, after changing the values like names in the file `preamble/mydefinitions.tex`.

Abstract: There should be a single paragraph of about 250 words, which concisely summarizes the entire proposal, written in the file `preamble/abstract.tex`.

Acknowledgments, Abbreviations, Glossary, Dedication preamble pages are optional and can be used at the author's discretion.

The main text of the proposal should be stored in the "SeniorThesis.tex" document. The following descriptions are sections that must be included in the thesis document.

Bibliography: The bibliography should include all references cited in the text (as [**?**]) and it should not include references that have not been cited. ACM referencing style should be used when preparing the bibliography. We recommend using BibTeX or BibLaTeX and using the file `preamble/bibliography.bib`.

Chapter 1

Introduction

1.1 Motivation

It is normal to have the situation that, when people cannot find the right music track for certain emotional states.

1.1.1 Human Emotion and Music Emotion

Humans like music for its capability to express emotions. The study has shown that music can evoke human emotions and influence human moods [18]. In the experiments mentioned by Przybysz, some sad or happy music can cause human physical reactions such as muscle tension and hormone release. This is one of the reasons for the human to have certain emotions toward music. However, sometimes music emotion and human emotion are separate. Human and music may carry the same emotion, for example, sadness, but human sadness and music sadness may be different. The human may be seeking for temporary separation from the real world. Context outside of music could also arouse human emotions. For example, when people are at the funeral, they are meant to be sad. When the lyrics or the music title are sad, people would also feel sad.

As music is so important to humanity and the internet became the most important tool in society nowadays, music service companies started to provide people access to music content and convenient service bundles. Based on different user demands, different music services have different strategies. For example, some users want to access a large number of music tracks with a low cost, some users want to discover new music content in a smart way [17]. Therefore, Spotify provides music stream services, and users have the option to upgrade to Spotify premium. With Spotify premium, users can skip songs, make tracks offline with users options, and more. With a good music service, users use music playlist published by other users to play at coffee shops, users listen to music when exercising, and users play music at weddings.

1.1.2 Easy Music Discover

With a large amount of music content users can access, it would be hard for users to pick the right song at the right moment, expand their playlists, or listen back to the song they saved previously [17]. Therefore, some music service providers developed music recommendation systems for users to find music content smartly. Spotify's Weekly Discovery is a good example for users to expand their playlist based on the music they saved previously. Generally speaking, the music recommendation system we can find on market works in two ways. The first way if to find new music content based on users' previous interested music. The second way is to find users who have similar actions and music preferences and cross recommend songs to each other. Since different users would not possibly have the same music playlist, it would be efficient to suggest new songs to different users.

These music recommendation services are accurate, users may enjoy it, but this system has a limitation. It only digs deeper for user interests, but it does not expand them. The reason is that the information for music recommender systems to use only contains user interaction toward music tracks. The system can classify and rank music tracks that people like or dislike, but they cannot get information about the user's emotional state.

1.1.3 User Current Emotion State

As mentioned before, the relationship between music emotions and human emotions are complicated. Not only some music can evoke human emotions, in some situations, but human emotions can also be evoked by the surroundings, lyrics, and humans may have their emotions and they are seeking similar emotions in music. In situations when people are sad, they want some corresponding sad music tracks to help them relieve mental painfulness. The recommendation system will recommend a list of songs based on user previously saved music. However, users would not only have sad music in their playlist, but they also have happy music. It would be hard for users to locate the songs they want. Although some music services have a sad music classification, users still have to explore more for the music that they want, because the scene defined by the lyrics may be different.

The important part here is to recognize the emotional state of the users. If music service recommends music content based on user current emotion state, the result would be more accurate in terms of user current demand. Therefore, it is better to let users express themselves. Let users define their emotions states themselves. This way, users can get music recommendations more accurate because users know their requirements the best. This solution can also provide a wider range of recommendations because it does not depend on user interaction history. Every music content suggested by this theory would be potentially new from users' past preferences.

1.2 Current State of the Art

There are various music services, such as Spotify, Amazon Prime Music, Pandora, and YouTube Music, that people like to use. The most popular music service on the market would be Spotify. According to MIDiA [15], a company analyzes media and technology, Spotify owns 83 million subscribers and 36% of the market share [16] at H1 of 2018.

1.2.1 Spotify

Spotify has different clients on different platforms. For example, there are mobile clients, desktop clients, TV clients, Web player, and more. Users can choose freely between different platforms. In each Spotify platform, the user can choose their preferred music content to play [24]. With smart home technologies got more and more popular, Spotify added functions that can control playback devices on different platforms. For example, users can control Google Home to play music from Spotify by smartphones. Besides the basic music player function, users can also add other users as friends to get other people's music content.

Spotify also has a strong music recommendation system to help people get more new music that they like. Users can get recommendations from Discover Weekly, Daily Remix, and song radios. There are three main techniques Spotify used to get recommendations for the users, which are Collaborative Filtering, Natural Language Processing, and Audio Metadata Modeling. In other words, Spotify utilized information from user interactive history and compares across different users, lyrics semantics, music patterns, and audio metadata, to get similar music contents.

1.2.2 YouTube Music

YouTube Music is another popular music service on the internet nowadays. YouTube was famous for its video services. YouTube users can upload videos they produced. Users who want to enjoy personalized services have to register user accounts. Based on user interests, YouTube would recommend videos on their homepage, and under each video that been played by the users, several related videos would be shown for users to pick. Under each video, users can write comments to share their thoughts. Video providers can also interact with other users. Video producer not only can gain popularity on YouTube but also money from YouTube because YouTube let companies put advertisements in the video contents. This is one of the reasons people are willing to upload videos to YouTube.

YouTube got more and more popular in recent years, they have announced that they have 100 hours of videos uploaded to them every hour in the past [14]. According to a user insight report in 2018, about 47% of users listen to music on-demand at Youtube [3]. With such a good user foundation, music producers and companies started to release music videos on YouTube. Based on a large number of licensed music videos, YouTube started music streaming services. With a good video recommendation system, YouTube combined content recommendation and social media functions. Similar users are easier to get each other's suggestions.

1.2.3 Netease Cloud Music

Netease Cloud Music is a Chinese music service provider. Most of the music license is held by Tencent Music at the time. As a new music service provider, to compete with Tencent Music, Netease Cloud Music developed stratify to develop strong music recommendation systems and social media platforms.

They think the essence of music service is to help users find the music they like efficiently. When users find the music they like, they would usually want to express themselves. Therefore, Netease Cloud Music provides comment sections for users to write their feelings and stories. Users can also share music between friends and other social media platforms. Because of the lack of music license, Netease Cloud Music encourages small and new producers to upload music to their platform. They also depend more on music recommendation between similar users. Due to Netease Cloud Music's good reputation among users, they got a great number of fundings recently [21].

1.3 Goals of the Project

A new system was proposed to match descriptive paragraphs with music tracks. This way, users can express their emotions whenever it is necessary, and a complex mechanism to detect the user's current emotion state can be avoided. Besides simple emotion detects, they can also expand their music interest by exploring more random but relevant music contents.

1.3.1 Scenes

Multiple scenes would be fit for the project. For example, a person breaks up with his or her girlfriend or boyfriend would be very sad at the moment. People would listen to sad songs when that happens. To find the right music tracks to identify their exact feelings, the person who is experiencing negative emotions can write literature paragraphs. These paragraphs would contain the sentiment of that person, and the story behind the negative emotions. It would be the same for people experiencing positive emotions, such as fall in love with someone.

The scene does not limit to self-generated emotions. For example, a person who watches a film or television could find himself or herself related to the work. The person would listen to related music tracks to express appreciation. Or the filmmakers want to find relevant music that matches the scene of the film, even contains matching metaphor. They can use a description of the film or the film scripts as the descriptive paragraph. By mapping the key factors of the paragraph with the music key factors, the relevant music tracks would be identified.

1.3.2 General Workflow

Before the user uses the service, pre-processing is required to get extra music contexts. The lyrics of the music track would be analyzed, and they would be summarized into keywords. And the result would be stored at the server database.

After the pre-processing is done, this system would firstly take user inputs. However, users are lazy, it always requires more to motivate people to generate sufficient inputs. Therefore, the format of the input has to be closely related to potential emotion breakouts. These inputs can be a diary, personal blog, video script, or description of the video content.

After getting a sufficient amount of user inputs, the system would analyze the inputs, breaks them down into pieces and summarize them into keywords. The sentiment of the text would also be analyzed. This extracted information would be used to compare with the language model in the music dataset. Specifically, the similarity would be checked between user description summarization and music fingerprints.

Finally, the system will return a list of relevant songs to users. The general workflow of the system shown as figure 1.1.

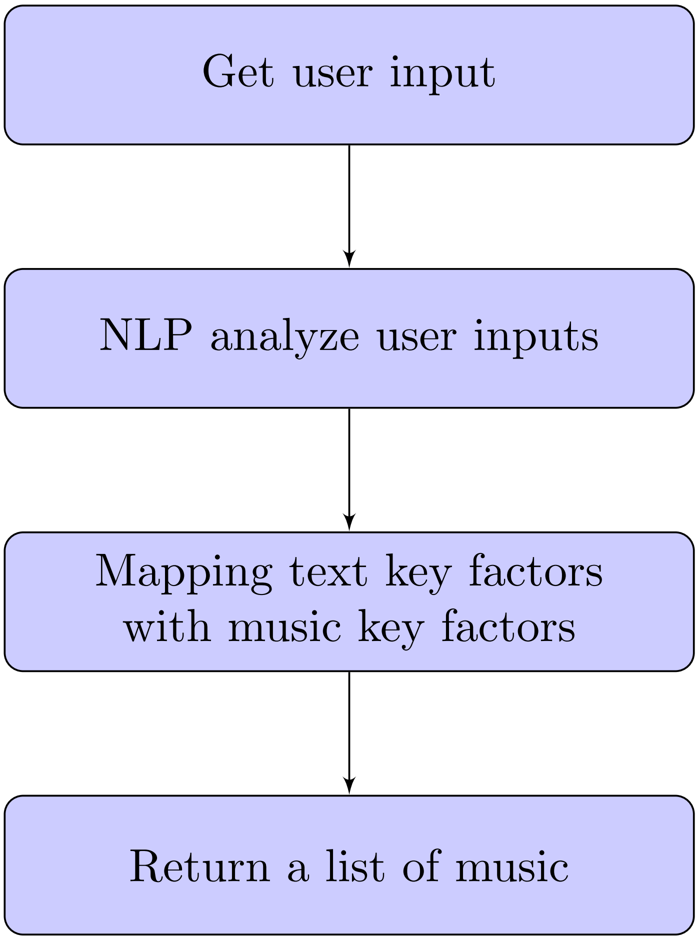

Figure 1.1: General Workflow

1.4 Thesis Outline

In this project paper, five sections will be used to introduce and analyze the proposed music recommendation system.

- The introduction will be mainly talking about the motivation of the project. Popular music services will also be introduced and analyzed. Besides motivation and popular music services, the expected goal of the music recommendation system and the general process steps will also be explained.

- The related work section will be mainly introducing similar studies. Each study will be introduced and analyzed toward their progress and limitations.
- The method section will be mainly about how is the system being build. Packages and tools used to build the system will also be introduced.
- The experiment section will be mainly talking about the evaluation of the system. The evaluation process and final results will be introduced and analyzed.
- Finally, the key findings and general ideas will be wrapped up in the conclusion section.

Chapter 2

Related Work

As a popular entertainment business, music service was always been a popular topic in the industry. Both the recommendation system and music emotions have many existing research projects.

2.1 Spotify Recommendation System

A recommendation system or recommender system is an algorithm that can suggest user-preferred items to a user by giving score or probability to items [20]. In the music recommendation system, the item to suggest would be music.

Take Spotify as an example, it has three parts in their recommendation system, which are Collaborative Filtering, Natural Language Processing, and Audio Metadata Modeling.

2.1.1 Collaborative Filtering

Collaborative Filtering [23] is an algorithm that recommends new content between similar users or items. The algorithm predicts user preferences by learning from users' past ratings for items [8].

Item-based

The first way to do collaborative filtering is to predict the relationship between items. Therefore, calculate item similarity is the first step. The common ways of calculating similarity for item i and j, sim(i,j), are cosine similarity, Pearson correlation, adjusted cosine, and conditional probability.

Let the targeted users be set U, $r_{u,i}$ and $r_{u,j}$ be user's rating over item i and item j. The cosine similarity can be calculated by equation (2.1).

$$\mathrm{sim}(i,j) = \cos(\mathbf{i},\mathbf{j}) = \frac{\mathbf{i}\cdot\mathbf{j}}{\|i\| * \|j\|} = \frac{\sum_{u\in U} r_{u,i} r_{u,j}}{\sqrt{\sum_{u\in U} r_{u,i}^2}\sqrt{\sum_{u\in U} r_{u,j}^2}} \tag{2.1}$$

Let $\bar{r}_u$ be the average rating for user u. The adjusted cosine can be calculated by equation (2.2) [1].

$$\operatorname{sim}(i,j) = \frac{\sum_{u \in U} \left(r_{u,i} - \bar{r}_u\right)\left(r_{u,j} - \bar{r}_u\right)}{\sqrt{\sum_{u \in U} \left(r_{u,i} - \bar{r}_u\right)^2}\sqrt{\sum_{u \in U} \left(r_{u,j} - \bar{r}_u\right)^2}} \tag{2.2}$$

Let $\bar{r}_i$ be the average rating for item i. The Pearson correlation can be calculated by equation (2.3). One good thing about correlation similarity is that it can calculate how close are the items related [8].

$$\operatorname{sim}(i,j) = \frac{\operatorname{Cov}(i,j)}{\sigma_i \sigma_j} = \frac{\sum_{u \in U} \left(r_{u,i} - \bar{r}_i\right)\left(r_{u,j} - \bar{r}_j\right)}{\sqrt{\sum_{u \in U} \left(r_{u,i} - \bar{r}_i\right)^2}\sqrt{\sum_{u \in U} \left(r_{u,j} - \bar{r}_j\right)^2}} \tag{2.3}$$

Finally, the conditional probability can be calculated by equation (2.4).

$$\operatorname{sim}(i,j) = P(j|i) = \frac{f(i \cap j)}{f(i)} \tag{2.4}$$

After obtained the similarity between items, a value of the item for the targeted user can be predicted. It means to obtain how the user rates similar items for specific item i, the record we obtained could is the history records.

User-Based

Another way to do collaborative filtering is to find the users that are similar. After knowing the similar users, and the item similarity scare are obtained, the recommendations can be predicted.

Take Spotify as an example, after collecting the user interaction, for instance, the playlist, whether liked the song, a unique coordinate will be generated for each user. The collected fields are the dimensions. The item-based suggestion will be coming from analyzing the user playlists, and the user-based suggestion will be coming from comparing different users. Once the similar users are identified, the algorithm will simply recommend contents that one user has and others do not.

2.1.2 Natural Language Processing

Natural language processing is the study of the interaction between computers and human languages. Human language can be speeches or writings. This study has many sub-fields, for example, speech recognition, text summarization, and text generation. By using natural language processing techniques, the meaning of lyrics can be classified, which can be further used to identify the music genre. The sentiment of the lyrics can also be analyzed, which can be used to identify the mood of the music tracks.

Text Sentiment Classification

Text classification is a very important sub-field of natural language processing, many real-life applications are depending on text classification. For example, web search, document classification, and information ranking [11]. The classification can be based on different aspects, in music recommendation examples, sentiment can be a feature. Sentiment analysis is the contextual mining of text emotions, for instance, positive or negative. Happiness, joyful, and compliment emotions can be categorized as positive emotions. Sad, angry, and guilt can be categorized as negative emotions. The general text classification workflow is shown in figure 2.1.

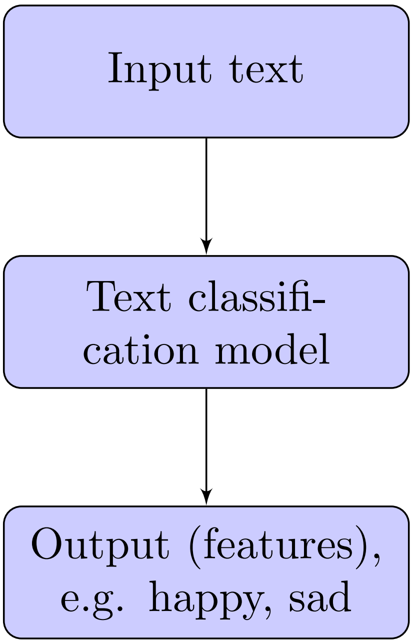

Figure 2.1: Text Classification Workflow

The first step to classify text is to treat text as a bag of words [6]. Bag of words is a model in natural language processing that treats texts as unordered words, and the frequency of each word would be recorded. This process is usually conducted by tokenization, and followed by stemming and lemmatization. Stemming and lemmatization can delete the additional part of the word, remain only the stem part. For example, after stemming and lemmatization, the word "interesting" would become "interest". After stemming and lemmatization, stop words, which are words that have no actual meaning in a sentence, will be removed from the bag of words. An example of a stop word can be "and".

One way to classify texts is a Naive Bayes algorithm. Naive Bayes is based on Bayes' probability theorem [19]. In Bayes' probability theorem, for event A and B, the conditional probability of A given B is equation (2.5).

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} \tag{2.5}$$

Multinomial naive Bayes is one of the ways to use the Naive Bayes algorithm to classify texts. It uses the term frequency, $tf(t, d)$, where t is term, and d is the active document. Let $tf(t, d)$ be the frequency of term t in document d, and n_d be the total number of terms in document. As the equation (2.6) shows.

$$\text{term frequency } = \frac{tf(t, d)}{n_d} \tag{2.6}$$

Spotify uses textual information from the internet to find key terms for music content. The terms will be used to build vectors for each song. A similar technique used in collaborative filtering will be applied to find similar content. This technique has a flaw that the textual information gets from the internet is too broad. It will be better if the information is from the comments under each song directly. That way, each song will have a direct correlation to the key terms. Except for comments, user search history could also be used as part of the text [7].

2.1.3 Audio Metadata Modeling

Audio metadata is the information for the audio files, it usually contains names for artist, album, title, genre, track number, and more [4]. After sampling the audio, more information can be included in audio metadata, for example, loudness, tempo, and more. This information can be used to predict potential music or human emotions. Tempo represents the speed of the audio, it is related to human emotions according to studies. Fast tempo and emotions like joy and fear have correlations, slow tempo and emotions like sadness and tenderness have correlations [12].

Spotify has a database that contains the audio metadata. These aspects can be used to build vector representations. By applying a similar technique in collaborative filtering, similar content will be identified. Bogdanov et al. proposed a way of classifying music by using audio metadata [5]. They asked participants to choose their preferred music and categorized them based on their preferences. They then modeled the audio files and classified them. After that, they correlated the classified features with predefined categories. By using this model, they can suggest songs based on audio metadata.

2.2 Content-Based Music Information Retrieval

Besides collaborative filtering, content-based music information retrieval is another effective way of recommending. This method collects information that describes the music, and use the information to suggest new content. Audio metadata can be one of the sources for the information.

One application is to collect and classify audio signal features, then apply them to match based on different situations. In general, there are two ways of matching, which are query by example, and query by humming [13]. Query by example means getting the audio signal as input and return music metadata as output. This query method works on the specific music track, but not the variation tracks, for example, a cover version. Query by humming makes it up by getting melody as input and returns similar tracks. This application is more of a searching or matching algorithm. The limitation of this recommendation method is as I mentioned previously, lack of novelty [13]. It only expands on user saved songs, and it cannot predict user situation.

The other application is to combine the information which describes the music with user preferences. The user preference does not mean the user ratings, content-based music recommendation does not rely on user ratings [8]. However, categorizing music content based on descriptive information requires experts of the fields to set rules. A similarity check is also required for the content-based suggestion. The way of calculating the similarity between two items is through calculating their distance. The common ways of calculating such distance are Euclidean distance (2.7), Manhattan distance (2.8), vector cosine distance(2.1), Mahalanobis distance (2.9), and Chebyshev distance (2.10).

$$d = |x - y| = \sqrt{\sum_{i=1}^{n} |x_i - y_i|^2} \tag{2.7}$$

Euclidean distance measures the straight line between two points in N-dimensional space [2].

$$d(x, y) = \sum_{i=1}^{n} |x_i - y_i| \tag{2.8}$$

Manhattan distance is also a measurement of the straight line between two points in N-dimensional space. Just imagine the street blocks in Manhattan, there two paths from one intersection to the other one. However, the true distance between them is a straight line cut through the buildings [9].

$$d(x, y) = \sqrt{(x - y)^T S^{-1} (x - y)} \tag{2.9}$$

Mahalanobis distance is a measurement of the straight line distance between two points in N-dimensional space. Mahalanobis distance handles better with correlated points(≥ 2). In situations that there are more than 2 points that are correlated, the axes of the points are no longer orthogonal, and they cannot be plotted in 3-dimensional space [22].

$$d(x, y) = \max_{i=1...n} |x_i - y_i| \tag{2.10}$$

Chebyshev distance, also called maximum value distance, is similar to Euclidean distance and Manhattan distance. It calculates the straight line distance based on point coordinates.

There are limitations for this application when new users entered the system, they can't get the right suggestion because the system needs time to adjust to the user preferences. This situation is also called the cold-start problem [8]. Another issue is feature extraction. It is hard to extract high-level descriptions such as mood. These high-level descriptions are very important in accurate and personalized recommendations.

2.3 Contextual Music Retrieval

Contextual music information is any information, other than music content information, that describes the music content itself. Contextual music suggestion means to recommend music based on the user's actual situation, for example, emotional state. This method takes three sources of information, environment-related context, user-related context, and multimedia context [13]. Environment-related information can be the season, temperature, time, and weather because all these contexts can influence human emotions. User-related information can be user activities, such as exercising and driving, user profiles, such as social network statements, and emotional state. Multimedia can be text and image relevant to the music. This information can be used to profile and predict the user's actual situation. After the necessary information is gathered, similarity predicts, classification, and other techniques can be used to predict user preferences.

Another way of achieving contextual music retrieval is proposed by B. Han et al [10]. They proposed an emotion state transition model which models music-evoked human emotions, and content-based music recommendation ontology to profile user preference and context. The system provides music by mapping high-dimensional music features with the emotion state transition model. According to the paper, this method achieved 67.54% overall accuracy. The limitation of this method is that it needs a large amount of data, and the human emotional state can be changed based on the social environment. Once the social context is changed, more data would be required to get higher accuracy.

Recommendation based on contextual music retrieval is still new, and it has great potential to contribute to the music recommendation system.

Chapter 3

Method of Approach

This chapter should answer the "how" question - how did you complete your project, including the overall design of your study, details of the algorithms and tools you have used, etc. Use technical diagrams, equations, algorithms, and paragraphs of text to describe the research that you have completed. Be sure to number all figures and tables and to explicitly refer to them in your text.

Chapter 4

Experimental Results

This chapter should describe your experimental set up and evaluation. It should also produce and describe the results of your study. Possible section titles are given below.

4.1 Experimental Design

4.2 Evaluation

4.3 Threats to Validity

Chapter 5

Discussion and Future Work

This is the conclusion. You might want to leave it unnumbered, as it is now. If you want to number it, treat it like any other chapter.

This chapter usually contains the following items, although not necessarily in this order or sectioned this way in particular.

5.1 Summary of Results

A discussion of the significance of the results and a review of claims and contributions.

5.2 Future Work

5.3 Conclusion

Bibliography

[1] Adjusted Cosine Similarity.

[2] 'n'-dimensional euclidean distance.

[3] IFPI releases 2018 Music Consumer Insight Report, Oct 2018.

[4] Metadata in digital audio files – what it is, where it is and how to tidy it up., Jan 2018.

[5] Bogdanov, D., Haro, M., Fuhrmann, F., Xambó, A., Gómez, E., and Herrera, P. Semantic audio content-based music recommendation and visualization based on user preference examples. *Information Processing & Management 49*, 1 (2013), 13–33.

[6] Brilis, S., Gkatzou, E., Koursoumis, A., Talvis, K., Kermanidis, K. L., and Karydis, I. Mood classification using lyrics and audio: A case-study in Greek music. In *IFIP Advances in Information and Communication Technology* (2012).

[7] Bu, J., Tan, S., Chen, C., Wang, C., Wu, H., Zhang, L., and He, X. Music Recommendation by Unified Hypergraph: Combining Social Media Information and Music Content. In *Proceedings of the 18th ACM international conference on Multimedia* (2010).

[8] Celma, Ò. *Music Recommendation and Discovery.* 2010.

[9] Craw, S. Manhattan distance, Jan 1970.

[10] Han, B. J., Rho, S., Jun, S., and Hwang, E. Music emotion classification and context-based music recommendation. *Multimedia Tools and Applications* (2010).

[11] Joulin, A., Grave, E., Bojanowski, P., and Mikolov, T. Bag of tricks for efficient text classification. *arXiv preprint arXiv:1607.01759* (2016).

[12] Kamenetsky, S. B., Hill, D. S., and Trehub, S. E. Effect of tempo and dynamics on the perception of emotion in music. *Psychology of Music 25*, 2 (1997), 149–160.

[13] Kaminskas, M., and Ricci, F. Contextual music information retrieval and recommendation: State of the art and challenges, 2012.

[14] Liikkanen, L. A., and Salovaara, A. Music on YouTube: User engagement with traditional, user-appropriated and derivative videos. *Computers in Human Behavior* (2015).

[15] MIDiA-Research. About Us, 2019.

[16] Mulligan, M. Mid-Year 2018 Streaming Market Shares, Sep 2018.

[17] Palaniappan, V. Spotify Product Analysis, Jan 2018.

[18] Przybysz, P. Music and emotions. *Avant* (2013).

[19] Raschka, S. Naive bayes and text classification I - introduction and theory. *CoRR abs/1410.5329* (2014).

[20] Ricci, F., Rokach, L., and Shapira, B. Introduction to recommender systems handbook. In *Recommender systems handbook*. Springer, 2011, pp. 1–35.

[21] Russell, J. China's NetEase raises $600M for its music streaming business, Nov 2018.

[22] Stephanie. Mahalanobis distance: Simple definition, examples, Aug 2018.

[23] Su, X., and Khoshgoftaar, T. M. A Survey of Collaborative Filtering Techniques. *Advances in Artificial Intelligence* (2009).

[24] Zhang, B., Kreitz, G., Isaksson, M., Ubillos, J., Urdaneta, G., Pouwelse, J. A., and Epema, D. Understanding user behavior in Spotify. In *Proceedings - IEEE INFOCOM* (2013).