CONTENT-BASED
MUSIC RECOMMENDATION:
A COMPARATIVE ANALYSIS OF FEATURES

NURUL AFEEQA BINTI MOHAMED NOR AZIZI

UNIVERSITI KEBANGSAAN MALAYSIA

CONTENT-BASED MUSIC RECOMMENDATION SYSTEM:
A COMPARATIVE ANALYSIS OF FEATURES

NURUL AFFEQA BINTI MOHAMED NOR AZIZI

PROJECT SUBMITTED IN PARTIAL FULFILMENT FOR THE DEGREE OF
MASTER OF DATA SCIENCE

FACULTY OF INFORMATION SCIENCE AND TECHNOLOGY
UNIVERSITI KEBANGSAAN MALAYSIA
BANGI

2024

CONTENT-BASED MUSIC RECOMMENDATION SYSTEM:
A COMPARATIVE ANALYSIS OF FEATURES

NURUL AFFEQA BINTI MOHAMED NOR AZIZI

PROJEK YANG DIKEMUKAKAN UNTUK MEMENUHI SEBAHAGIAN
DARIPADA SYARAT MEMPEROLEH IJAZAH SARJANA SAINS DATA

FAKULTI TEKNOLOGI DAN SAINS MAKLUMAT
UNIVERSITI KEBANGSAAN MALAYSIA
BANGI

2024

DECLARATION

I, Nurul Afeeqa binti Mohamed Nor Azizi, declare that this thesis entitled "Content-Based Music Recommendation System: A Comparative Analysis of Features" is the result of my work and has not been submitted for any other degree in any other university. All sources consulted and cited in this thesis have been duly acknowledged.

24 July 2024 NURUL AFEEQA BINTI MOHAMED NOR AZIZI
P127179

ACKNOWLEDGEMENT

First and foremost, I would like to praise Allah the Almighty, the Most Gracious, and the Most Merciful for His blessings throughout my Master's journey and in completing this dissertation.

I want to express my sincere gratitude to my research supervisor, Professor Dr Shahrul Azman bin Mohd Noah, for his invaluable guidance throughout this dissertation. I am truly fortunate to have been mentored by him, whose extensive knowledge significantly contributed to its completion. His meticulous attention to detail and unwavering support significantly impacted the quality of this dissertation.

I am also profoundly grateful to my classmates for their unwavering support and guidance. Their motivation and assistance were crucial in overcoming challenges in completing this thesis. Together, we have achieved a significant milestone.

ABSTRAK

Sistem rekomendasi ialah sistem penapisan maklumat yang bertujuan untuk menjangkakan penilaian yang ditetapkan oleh pengguna terhadap sesuatu produk. Ia adalah algoritma yang mencadangkan produk relevan kepada pengguna. Sistem rekomendasi boleh diklasifikasikan kepada penapisan kolaboratif, berasaskan kandungan, dan hibrid. Muzik adalah antara domain di mana sistem rekomendasi memainkan peranan penting. Sistem rekomendasi muzik dapat membantu pengguna menemui lagu yang selari dengan citarasa dan pilihan mereka, terutamanya apabila data muzik yang luas menghalang proses pencarian mereka. Disertasi ini menangani jurang ini dengan menilai prestasi model berasaskan kandungan yang dibangunkan melalui analisis perbandingan yang meluas bagi ciri-ciri yang disebutkan. Set data yang digunakan dalam disertasi ini ialah versi yang diubahsuai daripada Million Song Dataset, set data pengguna The Echo Nest dan set data Last.fm. Metodologi yang komprehensif telah digunakan untuk mencipta profil pengguna berdasarkan sejarah muzik pengguna yasng signifikan dengan ambang tertentu. Beberapa teknik praproses penting telah dijalankan, termasuk menangani nilai yang hilang, keluarkan atribut yang tidak relevan, menjalankan Kekerapan Jangka-Kekerapan Dokumen Songsang pada tag dan genre, dan menggunakan penskalaan ciri min-maks untuk ciri-ciri lagu berangka. Selepas langkah teknik praproses selesai, ciri daripada kategori berbeza telah dipurata untuk mewakili pilihan pengguna. Seterusnnya, persamaan konsinus digunakan sebagai ukuran persamaan untuk mencadangkan lagu dengan ciri yang serupa dengan pilihan pengguna. Penilaian luar talian telah digunakan untuk menilai prestasi model rekomendasi, dan metrik prestasi yang terlibat ialah ketepatan, kebolehkesanan dan skor F1. Hasil kajian menunjukkan bahawa menggabungkan ciri-ciri lagu dengan tag lagu dan genre boleh meningkatkan model rekomendasi dengan ketara dari segi ketepatan, kebolehkesanan dan skor F1. Disertasi ini menyumbang kepada pandangan berharga tentang potensi mengintegrasikan pelbagai data untuk meningkatkan kualiti rekomendasi muzik berasaskan kandungan, dengan itu meningkatkan pengalaman penemuan muzik pengguna pada platform penstriman muzik.

ABSTRACT

A recommender system is an information filtering system that aims to anticipate the rating or preference that a user may assign to an item. It is an algorithm that suggests relevant items to users. Recommender systems can be classified into collaborative, content-based and hybrid filtering. Music is among the domains in which recommender systems have played a pivotal role. Music recommendation systems play a crucial role in helping users discover songs that align with their tastes and preferences, especially when a vast amount of music data hinders their search process. More research on content-based music recommendation systems incorporating song features, song tags, and genres is needed, and the possibility of developing an enhanced recommendation is often overlooked. This dissertation addresses this gap by evaluating the performance of a developed content-based model through an extensive comparative analysis of the mentioned features. The dataset used in this dissertation is a remodelled version of the Million Song Dataset, The Echo Nest user dataset and the Last.fm dataset. Comprehensive methodology was employed to create user profiles based on significant user listening history with specific thresholds. Several essential preprocessing techniques were applied, which include handling missing values, removing irrelevant attributes, performing Term Frequency-Inverse Document Frequency on tags and genre and applying min-max feature scaling for numerical song features. After data preprocessing steps were completed, features from distinct categories were averaged to represent user preferences. Next, cosine similarity was employed as the similarity measure to recommend songs with similar features to the user preferences. An offline evaluation approach was applied to assess the performance of the recommendation models, and the performance metrics involved are precision, recall and F1-score. The results reveal that combining song features with song tags and genres can significantly improve the recommendation model's precision, recall, and F1-score. This dissertation contributes valuable insights into the potential of integrating diverse data for enhancing the quality of content-based music recommendations, thus improving users' music discovery experience on music streaming platforms.

TABLE OF CONTENTS

DECLARATION			iii
ACKNOWLEDGEMENT			iv
ABSTRAK			v
ABSTRACT			vi
TABLE OF CONTENTS			vii
LIST OF TABLES			x
LIST OF ILLUSTRATIONS			xi
LIST OF ABBREVIATIONS			xii
CHAPTER I	**INTRODUCTION**		
1.1	Research Background		14
1.2	Problem Statement		17
1.3	Research Question		20
1.4	Research Objectives		20
1.5	Research Scope		20
1.6	Significance Of Study		21
1.7	Thesis Organisation		22
	1.7.1	Chapter I: Introduction	22
	1.7.2	Chapter II: Literature Review	23
	1.7.3	Chapter III: Methodology	23
	1.7.4	Chapter IV: Results and Analysis	24
	1.7.5	Chapter V: Conclusion and Future Work	25
CHAPTER II	**LITERATURE REVIEW**		
2.1	Introduction		26
2.2	Collaborative Filtering		27
	2.2.1	Memory-Based Collaborative Filtering	27
	2.2.2	Model-Based Collaborative Filtering	28
2.3	Content-Based Filtering		30
2.4	Hybrid Filtering		33
2.5	Music Recommendation Systems		34
	2.5.1	Users	34
	2.5.2	Item Providers	35
	2.5.3	Music Streaming Platforms	36

2.6 Overview Of Content-Based Music Recommendation System 37

2.6.1 Defining the Dataset Variables and Features 45
2.6.2 Similarity Measures 50
2.6.3 Feature Representations 52
2.6.4 Evaluation of Recommendation Systems 53

CHAPTER III METHODOLOGY

3.1 Introduction 54
3.2 Dataset 55
3.3 Feature Categories 58
3.4 Research Tool And Software 58
3.5 User Profiles Creation 59
3.6 Data Cleaning 59
3.7 Feature Engineering 60

3.7.1 Term Frequency-Inverse Document Frequency 60
3.7.2 Min-Max Feature Scaling 61

3.8 Exploratory Data Analysis 62

3.8.1 Data Visualisation 63

3.9 Computation of Average Features for User Profiles 65

3.9.1 Song Features 65
3.9.2 Song Tags and Genres 66
3.9.3 Song Features and Tags, Genres 66

3.10 Generation of Music Recommendations Using Cosine Similarity 66
3.11 Recommendations Evaluation 68

3.11.1 Precision at K 68
3.11.2 Recall at K 68
3.11.3 F1-Score 69

CHAPTER IV RESULTS AND ANALYSIS

4.1 Introduction 69
4.2 Summary of User Profiles 70
4.3 Music Recommendations and Cosine Similarity Scores 73
4.4 Performance Evaluation 76

CHAPTER V CONCLUSION AND FUTURE WORK

5.1 Introduction 79

5.2	Research Summary	79
5.3	Limitations and Future Work	80
REFERENCES		**82**

LIST OF TABLES

Table No.		Page
Table 2.1	Summary of relevant research work	42
Table 2.2	Speechiness values and their characteristics	47
Table 2.3	Examples of tags based on various categories	49
Table 3.1	Music Information dataset attributes	55
Table 3.2	User Listening History dataset attributes	56
Table 3.3	Sample of dataset	57
Table 3.4	Feature categories and examples	58
Table 4.1	User profiles based on song features – Category 1	71
Table 4.2	User profiles based on song tags and genres – Category 2	71
Table 4.3	Example of recommended songs and user listening history of User 1 –Category 1	74
Table 4.4	Example of recommended songs and user listening history of User 1 – Category 2	75
Table 4.5	Example of recommended songs and user listening history of User 1 – Category 3	76
Table 4.6	Performance evaluation of content-based music recommendation systems using different feature sets	77

LIST OF ILLUSTRATIONS

Figure No.		Page
Figure 2.1	Types of recommendation systems	26
Figure 3.1	Research methodology	54
Figure 3.2	Missing values analysis	60
Figure 3.3	Histogram of preprocessed continuous song features	64
Figure 3.4	Bar chart of top song tags and genre	65

LIST OF ABBREVIATIONS

AI	Artificial Intelligence
ANN	Artificial Neural Network
BPM	Beats Per Minute
CBF	Content-Based Filtering
CF	Collaborative Filtering
CNN	Convolutional Neural Network
CoRS	Conversational Music Recommender System
CP-Net	Conditional Preference Network
CRNN	Convolutional Recurrent Neural Network
CSV	Comma Separated Values
dB	Decibels
DSP	Digital Service Provider
DT	Decision Tree
EDA	Exploratory Data Analysis
GMM	Gaussian Mixture Models
IDF	Inverse Document Frequency
IR	Information Retrieval
KNN	K-Nearest Neighbours
MAE	Mean Absolute Error
MF	Matrix Factorisation
MFCC	Mel-Frequency Cepstral Coefficient
MIDI	Musical Instrument Digital Interface
MIR	Music Information Retrieval
MRS	Music Recommendation System
ms	millisecond
MSD	Million Song Dataset
MUSIC	Mellow, Unpretentious, Sophisticated, Intense and Contemporary

NLP	Natural Language Processing
SPM	Sequential Pattern Mining
SVD	Singular Value Decomposition
SVM	Support Vector Machine
TF	Term Frequency
TF-IDF	Term Frequency-Inverse Document Frequency

CHAPTER I

INTRODUCTION

1.1 RESEARCH BACKGROUND

Music is an integral part of life and exists in all cultures. A diverse range of musical styles and features can be encountered in an individual's daily life, where one forms an opinion on whether they enjoy the music they hear (Greenberg et al. 2015). Research has shown that listening to music can positively impact one's life by relieving stress and enhancing mood through dissociation from an unpleasant state of mind (H. Immanuel James et al. 2019). Historically, music was limited to physical formats such as vinyl, cassette tape, and compact disc, which not all music listeners can access (Bauer & Schedl 2019). As we reach the 2010s, music streaming platforms have been multiplying yearly in data, revenue, and users (Curry 2023). These platforms revolutionise the music industry with the current rapid technology of recommendation systems, making artists and their music more accessible online and allowing listeners to discover music according to their preferences (Dinnissen & Bauer 2022).

As the name implies, a recommender or recommendation system suggests items to users or predicts items the users might be interested in. "Items" are a general term that refers to a collection of features, sometimes called properties or attributes, which act as an indication of what the system aims to suggest to its users (Bauer & Schedl 2019). Items are often stored in a database table, where the characteristics or attributes describe each item. This system has various filtering engines or approaches, broadly categorised as collaborative, content-based, and hybrid filtering (a combination of collaborative and content-based filtering). By obtaining relevant information, these approaches provide personalised recommendations to the users, but each approach has different algorithm techniques to generate the recommendations (Pipilis 2023).

For instance, collaborative filtering (CF) focuses on collaborating among users with similar behaviour and recommending items to the following user based on their collective choices (Song et al. 2012). Memory-based filtering and model-based filtering are the two types into which CF is separated (Roy & Dutta 2022). Memory-based CF predicts item ratings by analysing the complete collection of past ratings. User ratings are typically arranged in a user-item matrix called the utility matrix. Similar interest users are grouped, and predictions for a new item are generated by identifying the nearest neighbours using many explicit user votes. The categories of memory-based CF include user-based and item-based approaches. The user-based approach finds users with similar tastes to the target user and then recommends items those similar users like, while the item-based approach identifies items that are similar to ones the user has already enjoyed and recommends them. In contrast, model-based CF represents user preferences through a set of rating scores and constructs a specialised prediction model by employing data mining and machine learning algorithms.

Meanwhile, content-based filtering (CBF) in recommendation systems utilises various item features to construct a user profile, allowing for personalised recommendations based on user preferences and interests (Bauer & Schedl 2019). These systems match a user profile's attributes with the items' characteristics or multimodal information to provide tailored recommendations (Afchar et al. 2022; Okada et al. 2022). This dissertation is about music recommendation systems (MRS), which employs a content-based approach utilising features that describe the song, such as acousticness, danceability, energy, instrumentalness, loudness, liveness, speechiness, tempo, valence, song tags and genres. By analysing the features of a user profile and comparing them to the overall features of different songs, content-based MRS helps determine the user's level of interest in each song, leading to personalised music recommendations (Afchar et al. 2022).

The effectiveness of a recommendation system in retrieving and presenting relevant recommendations depends on how accurately the user profile reflects the user's preferences. User profile typically consists of various information that includes (Pazzani & Billsus 2007):

1. Description of user's favourite items: This can predict the probability of the user being interested in a particular item, efficiently identifying the top number of items that most likely capture the user's interest.
2. User's past behaviour or interactions with the items: For example, in the music domain, a user's ability to listen, like or save a song to a playlist represents positive interaction with the item. Positive or negative interaction serves various purposes, such as displaying recently accessed songs, filtering out songs the user has already listened to or interacted with, and providing training data for machine learning algorithms to enhance the user model in recommendation systems.

Past research has also demonstrated that combining Artificial Intelligence (AI) and Information Retrieval (IR) can significantly improve recommendation systems. Both CF and CBF can benefit from AI and IR, but the specific techniques used might differ depending on the focus on user interactions or item content attributes (Lops et al. 2011; Timanshi Bhardwaj et al. 2021).

Instead of continuously requiring users to provide information about their preferences to the recommendation system, the system can learn about users' preferences independently. This is the context in which AI's machine learning application is applied (Dinnissen & Bauer 2022). Machine learning is a process that consistently learns and aims to classify new information items by leveraging the user's past behaviour, where items have been labelled as preferred (like) or unpreferred (dislike). A predictive model can be created through these labelled items and machine learning techniques, hence determining whether the target user may like or dislike the new information item.

Due to their practicality and success, recommendation systems are widely used in various domains (Timanshi Bhardwaj et al. 2021). For instance, recommendation systems in the e-commerce industry suggest products the user might be interested in purchasing. Furthermore, these systems recommend movies the user might want to watch in the film industry. Recommendation systems assist in providing content suggestions to users, aligning new items with their interests across these various domains. Their primary aim is to deliver information tailored to users' needs based on factors like users' preferred content and personal history. Consequently, a recommendation system is considered effective when users explore the recommended content and find it appealing (Timanshi Bhardwaj et al. 2021). Major corporations such as Netflix, Amazon, and YouTube are known to have recommendation systems within their platform, as this can ease and automate the process of suggesting new items to their user base. However, the music domain differs from other recommendation tasks in such a way that it has the following characteristics (Afchar et al. 2022):

1. Shorter consumption time: The duration of songs typically lasts only a few minutes, unlike reading books or watching movies.
2. Multimodal data: Music data comes in various forms, which include the audio, Musical Instrument Digital Interface (MIDI), textual metadata (artist names, song titles, genres, song tags, lyrics, and other information about the song), cover arts or album covers (image), and even music videos.
3. Contextual influence: Music preferences are strongly influenced by the listener's location, social situation, and mood.
4. Sequential consumption: Songs in a playlist or listening session are often consumed in order. This leads to focusing on sequential recommendation tasks, such as creating or updating playlists considering short and long-term user preferences.

1.2 PROBLEM STATEMENT

Studies on recommendation systems have drawn significant attention from academicians and researchers. However, most research on recommendation systems seems to focus on the movie domain. This is probably because movie datasets are widely available and easily accessible to researchers (Roy & Dutta 2022). Consequently, there is a gap concerning recommendation systems in the music domain (Dinnissen & Bauer 2022; Roy & Dutta 2022). Addressing the gap is crucial, especially in this evolving landscape of digital music consumption. The availability of more advanced internet technology has led music streaming subscription services to emerge and continuously gain traction among music listeners (Polaris Market Research 2023).

The statistics of Spotify show that Spotify has over 100 million songs, 551 million monthly active users and 220 million subscribers (Iqbal 2024). As the global music streaming market continues to experience growth in both its song library and user base, the continuous refinement of MRS is vital to ensure increased quality.

Traditional CF approaches widely used in recommendation systems often rely heavily on recommending similar items to users with similar behaviours and preferences, overlooking the unique characteristics of individual items as done in CBF. Besides, the same group of users may not necessarily share similar tastes. Developing innovative strategies and techniques is necessary, as they can improve the quality of MRS and ensure user satisfaction. Few researchers have proposed approaches to improving the recommendation quality in content-based music recommendation as follows:

1. Niyazov et al. (2021): Content-based MRS was developed to consider a song's acoustic features. Rather than using various song features such as danceability, energy, loudness, speechiness and many more, their research only contributed to understanding each song's acoustic profile. This made the recommendation system suggest songs centred only on acoustic similarity to the user's preferences. The research provided limited data on song features, which resulted in relatively low values of quality metrics (Niyazov et al. 2021).

2. Adiyansjah et al. (2019): Deep learning techniques were used to extract features from audio signals, and the role of music genres in the recommendation process was emphasised. Their study also suggests that future research could explore adding other song features, such as using a tempo gram to capture the local tempo at a specific time, to improve the MRS' accuracy (Adiyansjah et al. 2019).

3. Appalanidu et al. (2021): Focused on developing personalised MRS using supervised machine learning algorithms. Their research involved predicting a song's mood and suggesting songs based on the mood similarity scores by

analysing the song lyrics. This research emphasised the importance of considering genres to improve the MRS quality (Appalanidu et al. 2021).

While previous research has explored content-based music recommendation, combining various features to enhance recommendation performance has been largely unexplored. For example, Niyazov and colleagues (2021) focused solely on acoustic features, Adiyansjah and colleagues (2019) on audio signals and genre-based features, and Appalanidu and colleagues (2021) on lyrical features.

Therefore, this dissertation is determined to develop and analyse the impact of combining multiple features into a content-based recommendation system within the music domain. This dissertation has the potential to enhance music recommendation systems significantly. To achieve this, three types of feature categories were chosen which include song features, song tags and genres, and the combination of all features. Examples of feature categories are displayed in Table 3.4, and these features were selected based on their potential to capture different dimensions of a song's profile relevant to listeners' preferences. For instance, song features provide measurable audio characteristics of each unique song that vary from one another (Mcfee et al. 2012). Song tags and genres give information about the songs that allow categorisation based on common characteristics or themes. The difference between the two is that genre is a much more general category, while tags provide more specific information describing the song's mood and style (Mcfee et al. 2012).

The need for a comparative analysis arises from determining the effectiveness of these individual features and their combinations in improving the recommendation performance. This dissertation undertakes a comprehensive comparative analysis to understand how each feature contributes to the recommendation system. Combining song features, song tags, and genres is assumed to produce better recommendation performance than a single feature. This aligns with machine learning principles, where providing more relevant information or data to the algorithm can improve performance. The thorough comparative evaluation among these different categories is explained in the later chapters, providing insights into each feature set and their combinations in the context of MRS.

1.3 RESEARCH QUESTION

Based on the provided problem statement, the following is the research question:

1. How do various feature categories, including song features, tags and genres, and a combination of both, impact the evaluation performance of a content-based music recommendation system in terms of precision, recall, and F1-score?

1.4 RESEARCH OBJECTIVES

This dissertation aims to find the best-performing content-based music recommendation model for users that aligns with their distinct music interests and preferences. A comprehensive investigation of song features, song tags, and genres, as well as a combination of both, is conducted by analysing the precision, recall, and F1-score. Strategies and techniques contributing to accurate and personalised music recommendations are also identified. The following research objectives guide this dissertation:

1. To develop a content-based music recommendation model that incorporates three categories: song features, song tags and genres, and a combination of both, for more precise music recommendations.
2. To evaluate and compare the evaluation metrics across three categories of the developed content-based music recommendation models to determine which category provides the best recommendation performance.

1.5 RESEARCH SCOPE

This dissertation encompasses a remodelled version and subset of multiple datasets known as the Million Song Dataset (MSD), The Echo Nest user dataset and the Last.fm dataset. The MSD was created to analyse song features and metadata of a million popular songs (Bertin-Mahieux 2011; Mcfee et al. 2012). The Echo Nest user dataset consists of real anonymous users of multiple music streaming platforms, songs from the MSD, and the number of counts (play counts) the user has listened to the song. The Last.fm dataset depicts the song tags and genres of the MSD. These datasets, which are widely accessible, are often used by researchers who are particularly interested in exploring Music Information Retrieval (MIR) and developing algorithms intended for commercial-scale data (Bertin-Mahieux 2011; Mcfee et al. 2012).

The remodelled version of the datasets used for this dissertation is directly obtained and downloaded from Kaggle, an online website for public use, especially for machine learning practitioners and data scientists. The dataset has two Comma-separated Values (CSV) files titled 'Music Information' and 'User Listening History'. 'Music Information' provides detailed information about 50,683 songs with 21 attributes. 'User Listening History' contains 9,711,301 user interactions with songs, each corresponding to a unique combination of the track identifier attribute and highlighting the play counts made by various users.

This dissertation does not employ a traditional machine learning approach for training or testing but relies significantly on a similarity-based approach for recommendations. This demonstrates that music recommendations can still be generated meaningfully without the need for complex model training.

1.6 SIGNIFICANCE OF STUDY

Recommendation systems are technological advancements that aim to provide users with easier access to a vast amount of online information (Roy & Dutta 2022; Wang et al. 2022). Since personalised recommendations are widely applicable in diverse domains, constant improvements and a deeper understanding are needed to adapt seamlessly to the digital era. This ensures increased IR efficiency for users and minimises the effort required to find their most preferred items. This dissertation focuses on a content-based MRS, emphasising song features, song tags, and genres instead of traditional external music descriptors like artist name and song title.

Continuous improvements in MRS benefit not only the users but also present a significant opportunity for the music industry, including music streaming platforms, music labels, and artists. These ongoing developments allow a better understanding of listeners' preferences and requirements, contributing to the continuous evolution of user experience and the music industry. Research has demonstrated that developing a personalised MRS with outstanding performance correlates with user satisfaction, which is necessary for maintaining user loyalty on music streaming platforms (Afchar et al. 2022). This dissertation also aims to contribute substantially to developing recommendation systems, particularly in the music domain, as research on this domain is scarce (Dinnissen & Bauer 2022). It highlights the importance of extracting song features, song tags, and genres to improve the quality of MRS. Additionally, this dissertation compares and evaluates each feature category to determine its effectiveness in enhancing MRS.

1.7 THESIS ORGANISATION

This section on thesis organisation provides the readers with a clear understanding of the research overview and structure. This dissertation consists of five chapters, and below is a brief description of the chapters, their sections and sub-sections:

1.7.1 Chapter I: Introduction

The first chapter consists of seven sections, which cover the background of the research, the problem statement, the research question, the research aim and objectives, the research scope, the research significance and, lastly, this section, which is the thesis organisation.

The research background discusses music and its positive impact on people's lives. It also discusses the evolution of music consumption and the role of recommendation filtering techniques in developing music recommendation systems. The problem statement of this dissertation argues that there is not enough research on recommendations specifically in the music domain, even though music consumption is an evolving landscape in the digital era. It also argues that CF techniques are widely used in recommendation systems, overlooking the benefits of CBF techniques. The research question addresses the impact of various feature categories on the evaluation performance of a content-based MRS. Specifically, it investigates how different sets of features, such as song features, tags and genres, and a combination of both sets, influence the system's precision, recall, and F1-score.

This dissertation aims to develop the best-performing content-based model in the music domain, guided by several research objectives. The first research objective is to develop models incorporating three distinct categories (song features, tags and genre, and a combination of all features). The second research objective is to evaluate and compare the evaluation metrics (precision, recall, F1-score) of the models' recommendation performance. The research scope highlights the scope of the research dataset. This dissertation utilises comprehensive remodelled datasets from Kaggle, combining the MSD, The Echo Nest dataset and the Last.fm dataset. The research significance was also discussed, highlighting MRS' continuous refinement and how it can benefit the music industry ecosystem.

1.7.2 Chapter II: Literature Review

Chapter 2 contains six sections, mostly about defining, explaining, and discussing topics related to recommendation systems and their music domain.

Various recommendation types and their filtering techniques were defined. After noting the recommendation definitions, the roles of various stakeholders, such as users, item providers, and music streaming platforms, are discussed to understand music recommendation systems better.

This dissertation is concerned with developing a content-based MRS and evaluating its effectiveness. Thus, this chapter presents previous studies on MRS, highlighting the content-based approach and cosine similarity among other similarity measures. The dataset variables and features were also defined, as well as the feature representations on content-based MRS and the evaluation of recommendation systems.

1.7.3 Chapter III: Methodology

Chapter 3 consists of eleven sections. This chapter discusses the research dataset, tools and software, research methodologies, exploratory data analysis, and evaluation of recommendations for developing a content-based MRS.

To delve deeper, a table summarising the dataset's attributes and a sample of the dataset was provided. Next, this chapter also highlights the use and execution of Python codes for developing the content-based MRS.

The research methodology emphasises the importance of creating user profiles based on specific criteria and thresholds to optimise memory usage and computational efficiency for Python execution. A table illustrating the user profile is also displayed. Several essential preprocessing techniques are discussed and applied to ensure increased data quality. Detailed explanations and formulas for each feature engineering technique, such as term frequency-inverse document frequency and min-max feature scaling, are also provided. Exploratory data analysis was also performed to gain visualisation insights about the preprocessed data.

After explaining data preprocessing techniques, the next step is to explain the process of averaging features from different feature categories to represent user preferences. The next step involves generating song recommendations using cosine similarity. Lastly, the research methodology ends with discussing the evaluation of the recommendations using precision, recall, and F1-score metrics.

1.7.4 Chapter IV: Results and Analysis

Chapter 4 has only four sections focusing on the results and analysis of user profiles, cosine similarity scores of music recommendations and the evaluation metrics of the developed content-based MRS.

The results obtained from the recommendation models for three distinct feature categories, song features, tags and genre, and combinations of all features, are evaluated and compared. The results are summarised and presented in a table so readers can easily compare and understand the recommendation systems’ performance metrics. The results of this dissertation are also compared with other researchers’ content-based music recommendation models.

1.7.5 Chapter V: Conclusion and Future Work

The final chapter of this dissertation has three sections. This chapter delivers a complete summary of the dissertation based on the methodologies, results, and analysis. The following sections provide limitations and suggestions for future work.

CHAPTER II

LITERATURE REVIEW

2.1 INTRODUCTION

Chapter 2 delves into the concepts and terminology of different recommendation system types with varying filtering techniques, such as collaborative filtering, content-based filtering, and hybrid filtering approach, as shown in Figure 2.1. Next, the roles of various stakeholders that make up the music industry ecosystem are also discussed. Since this dissertation is mainly about developing content-based MRS, the overview of previous research related to this approach is presented, including the specific arguments and ideas.

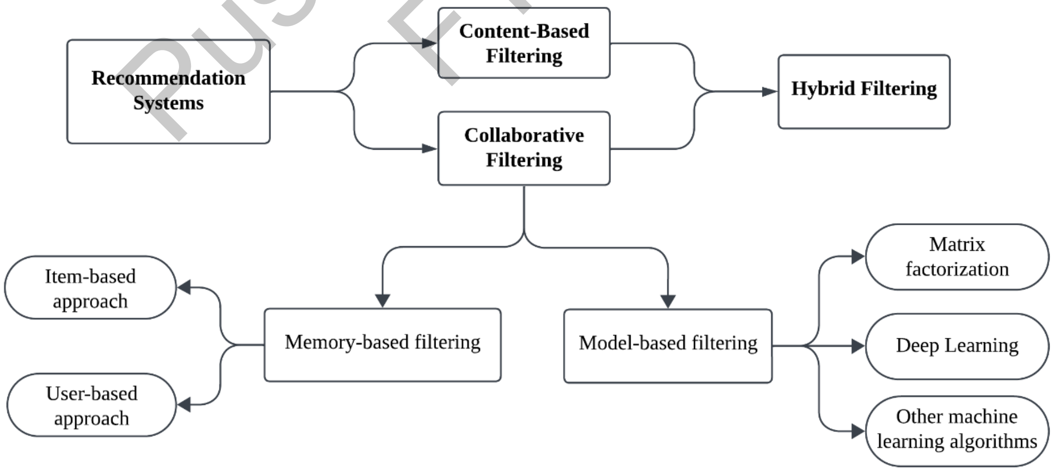

Figure 2.1 Types of recommendation systems

Source: Deepjyoti Roy & Mala Dutta 2022

2.2 COLLABORATIVE FILTERING

Collaborative recommendations apply CF which leverages the likings of similar users to recommend items since this recommendation assumes users with similar preferences have similar tastes (Mallannagari et al. 2020). Crucially, CF does not require any external information regarding item features or users to generate a recommendation, but they do require information on user-item interactions (Bauer & Schedl 2019; Roy & Dutta 2022). As illustrated in Figure 2.1, CF encompasses two main branches, model-based filtering and memory-based filtering, whereby each filtering approach has its own method for recommendation generation.

2.2.1 Memory-based collaborative filtering

Memory-based CF bases its recommendation system on how similar users or items are to one another (Nnadozie 2021). Construction of a utility matrix is the initial step followed by the identification of similar users and items, which is referred to as neighbourhood selection. Utility matrix, also known as a user-item matrix or preference matrix, is a structured representation of the interactions between users and items. Every cell in the matrix holds a user's feedback or rating regarding a certain item. This matrix serves as the foundation for CF algorithms, providing valuable information that represents users' preferences for items (Kim et al. 2023). In neighbourhood selection, a group of users is identified as a neighbourhood whose preferences closely align with a target user, and recommendations are derived from items favoured by this user group (Roy & Dutta 2022).

By establishing neighbourhoods of users or items with comparable preferences, the algorithms can accurately predict a user's preference for an item by considering similar users' preferences (Koren & Bell 2011). However, it is essential to note that neighbourhood selection is primarily used in memory-based CF, specifically with item-based and user-based approaches. For instance, the item-based approach in memory-based CF establishes recommendations by identifying users who liked a particular item and recommending other items that were liked by those users or similar users. This approach is item-centric and generates recommendations for other items (Nnadozie 2021). In contrast, the user-based approach in memory-based CF identifies users similar to a given user based on patterns of ratings or usage and recommends items liked by those similar users (Nnadozie 2021).

Once the neighbourhood of the user or items has been identified, similarity measures are carried out to quantify the similarity between users or items (Koren 2009). Various similarity measures include Cosine similarity, Euclidean distance, Pearson correlation coefficient, Jaccard similarity, and others (Roy & Dutta 2022). While each measure emphasises different aspects, their shared objective lies in determining the measure of similarity between users or items based on their interactions in the utility matrix (Han et al. 2011).

2.2.2 Model-based collaborative filtering

Model-based CF, on the other hand, involves building predictive models based on the observed user-item interactions and does not primarily rely on neighbourhood selection (Roy & Dutta 2022). Model-based CF utilises matrix factorisation (MF), deep learning and other machine learning algorithms to learn underlying relationships and predict the user's rating of unrated items (Nnadozie 2021; Roy & Dutta 2022). Past research has demonstrated that model-based filtering tends to have better accuracy recommendation performance than memory-based filtering (Zhang et al. 2014).

Several vital components or processes occur in model-based CF which include MF, K-Means clustering and deep learning. To establish recommendations, the technique of latent factor models such as MF also known as singular value decomposition (SVD), aims to map both items and users into a shared latent factor space to explain ratings by defining both items and users based on factors derived from user feedback (Koren & Bell 2011). MF can be represented in mathematical terms as shown in Equation 2.1:

$$\hat{y}_{ui} = f(u, i \mid \boldsymbol{p}_{u'} \boldsymbol{q}_i) = \boldsymbol{p}_u^T \boldsymbol{q}_i = \sum_{j=1}^{K} p_{uk} q_{ik} \qquad \ldots(2.1)$$

The prediction score is symbolised as $\hat{y}_{ui}$, u is the latent vector for the user, whereas $\boldsymbol{q}_i$ is the latent vector for the item, and the dimension of the latent space is K. Users and items have hidden factors captured by vectors which are $\boldsymbol{p}_{u'}\boldsymbol{q}_i$. The strength of each factor in the user or item is reflected in whether the vector's values are positive or negative. To see how well the user and item match, the system multiples their factor vectors. A high score indicates that the user is interested in the item's characteristics and prefers the item (Nnadozie 2021).

Moving on to the next vital component in model-based CF, K-Means clustering in model-based CF functions by randomly selecting a number of K items as the initial cluster centres, and each item is then assigned to a cluster to minimise the distance between the item and the cluster centre (Pipilis 2023). Afterwards, the distance is calculated using item similarity and for each cluster, the mean is recalculated based on the current cluster members (Nnadozie 2021). This whole process of assigning items, recalculating means and adjusting distances continues until it reaches convergence. The model-based CF utilises an unsupervised learning model to determine similarities (Nnadozie 2021). To further enhance and optimise the accuracy of similarity calculations, an overlap factor is introduced as shown in Equation 2.2:

$$\text{Isim}'\ (u, v, j) = \frac{\text{Min}\left(I_u \cap I_v \cap c_j \mid \gamma\right)}{\gamma} \times \text{Isim}\ (u, v, j) \qquad \ldots(2.2)$$

The original similarity score between the users u and v in the item cluster c_j is represented by Isim′ (u, v, j). $\left(I_u \cap I_v \cap c_j \mid \gamma\right)$ is defined as the number of items rated by the same set of users, u and v in the item cluster c_j, where γ is the parameter that increases with the number of co-rated items and essentially gives more weight to similarities based on a larger number of shared ratings (Nnadozie 2021).

Significant innovation has occurred in recommendation systems, where deep learning is applied to generate personalised recommendations (Schedl 2019). Deep learning is the third vital component of model-based CF which is explained in this dissertation. Deep learning includes neural networks, such as artificial neural networks (ANN) and convolutional neural networks (CNN) (Nnadozie 2021). Traditional model-