Faculty of Engineering and Technology

Department of Electrical and Computer Engineering

ListenTo web Application with recommendation system

Prepared by

Qossay ZeinEddin **1180235**

Amro Abu Hashish **1180620**

Mohammad Raddad **1180812**

Supervised by:

Dr. Hanna Bullata

A graduation project submitted to the Department of Electrical and Computer Engineering in partial fulfilment of the requirements for the degree of B.Sc. in Computer Engineering

BIRZEIT UNIVERSITY

January 2023

Abstract

Music has been an integral part of human culture since ancient times. Over the years, it has evolved into a modern form of communication that resonates with the current generation. Music is not just an art form but also a tool that helps us express our emotions, be it happiness, sadness, love, anger, or motivation. It is a universal language that connects people from all walks of life and unites them through the shared experience of music.

In today's digital age, the internet has become a vast repository of music from different genres and artists. While this has made it easier for people to access music from around the world, it has also made it difficult to find the songs that appeal to individual preferences. With so much music available online, it can be overwhelming to navigate through the endless lists of songs and artists.

To address this challenge, we have created a platform that not only provides a vast collection of songs but also uses artificial intelligence to recognize the user's preferences and mood. Our platform is designed to provide a unique experience to both users and musicians. It offers a personalized space where users can listen to their favorite tracks and discover new artists that match their preferences. It also provides a platform for musicians to showcase their talent to a wider audience, they can book available theaters and venues for a music concert with the ability to sell tickets through the website

Our website ListenTo is designed to be user-friendly and intuitive, making it easy for users to navigate and find the songs they love. The platform uses machine learning to suggest songs based on the user's listening history and preferences. This means that the more a user listens to music on our platform, the better the system becomes at predicting the songs they will enjoy. Additionally, our platform suggests artists similar to those that the user already likes, making it easier to discover new music.

المستخلص

كانت الموسيقى جزءًا لا يتجزأ من الثقافة البشرية منذ العصور القديمة. على مر السنين، تطورت إلى شكل حديث من أشكال الاتصال الذي يتردد صداها مع الجيل الحالي. الموسيقى ليست مجرد شكل من أشكال الفن ولكنها أيضًا أداة تساعدنا في التعبير عن مشاعرنا، سواء كانت السعادة أو الحزن أو الحب أو الغضب أو الدافع. إنها لغة عالمية تربط الناس من جميع مناحي الحياة وتوحدهم من خلال التجربة المشتركة للموسيقى.

في العصر الرقمي الحالي، أصبح الإنترنت مستودعًا واسعًا للموسيقى من مختلف الأنواع والفنانين. في حين أن هذا سهّل على الأشخاص الوصول إلى الموسيقى من جميع أنحاء العالم، إلا أنه جعل من الصعب أيضًا العثور على الأغاني التي تناسب التفضيلات الفردية. مع توفر الكثير من الموسيقى عبر الإنترنت، قد يكون التنقل عبر قوائم لا تنتهي من الأغاني والفنانين أمرًا مربكًا.

لمواجهة هذا التحدي، أنشأنا منصة توفر مجموعة كبيرة من الأغاني، تستخدم أيضًا الذكاء الاصطناعي للتعرف على تفضيلات المستخدم ومزاجه. تم تصميم منصتنا لتوفير تجربة فريدة لكل من المستخدمين والموسيقيين. يوفر مساحة مخصصة حيث يمكن للمستخدمين الاستماع إلى مساراتهم المفضلة واكتشاف فنانين جدد يتوافقون مع تفضيلاتهم. كما أنه يوفر منصة للموسيقيين لعرض مواهبهم لجمهور أوسع، ويمكنهم حجز المسارح والمدرجات المتاحة لعمل حفل غنائي مع إمكانية بيع التذاكر من خلال الموقع.

تم تصميم الموقع الإلكتروني ليكون سهل الاستخدام وبديهيًا، مما يسهل على المستخدمين التنقل والعثور على الأغاني التي يحبونها. يستخدم النظام الأساسي التعلم الآلي لاقتراح الأغاني بناءً على سجل الاستماع وتفضيلات المستخدم. هذا يعني أنه كلما زاد استماع المستخدم إلى الموسيقى على منصتنا، أصبح النظام أفضل في التنبؤ بالأغاني التي سيستمتعون بها. بالإضافة إلى ذلك، تقترح منصتنا فنانين مشابهين لمن يحبهم المستخدم بالفعل، مما يسهل اكتشاف الموسيقى الجديدة.

Table of Contents

Acronyms and Abbreviations

ML	Machine Learning.
ANN	Artificial Neural Network.
SRS	System Requirement Specification
API	Application Programming Interface
OOP	Object Oriented Programming
SQL	Structured Query Language
UI	User Interface.
UX	User Experience
XD	Experience Design
XML	Extensible Mark-up Language
ORM	Object–relational mapping
AWS	Amazon Web Services
RS	Recommender System
DBS	The Davies-Bouldin Score

List of Figures

List of Tables

Chapter 1: Introduction

Technology usage in the music platform has become essential due to technology improvement and its entry in many fields. Many research papers talk about the importance of using web and machine learning in the music platforms.

1.1 Motivation

The problem studied in this thesis is related to difficulty of finding music or songs due to the big amount of the songs and it will help you to find your taste, the system will be learning about you and what you like and try to suggest best songs according to the machine learning output it will be done automatically

1.2 Research goals

The objectives of the project are:

i. Reduce time of searching on songs
ii. Build a ML best model that suits our use to learn about the user
iii. Training the model to detect a new Froud cases that can happen
iv. Build a Web site to manage the artists and member business logic with songs, parties, paying and interact.

1.3 Report Organization

This report discusses topics to the project of ListenTo and its implementation

Chapter 2: In this chapter we will give background on ListenTo that is use and we will view a related work and literature review to our project.

Chapter 3: In this chapter we will discuss the Programming language used and its Framework (technologies) to use it in our implementation of the project also the Backend database and system architecture.

Chapter 4: In this chapter we will design the system and define user requirements, system requirements, Use case diagrams and activity diagrams.

The Last chapter includes two sections. Section one provides a conclusion of what we have reached so far in our project. And finally, section two plans an outline for future work to be done at the end of ENCS530.

Chapter 2: Literature Review

In this chapter, we will review relevant project with some similar ideas for ListenTo and discuss the different between our project.

2.1 Spotify

Spotify is a music streaming service that offers a vast library of music and allows users to create and share playlists, and listen to radio stations. It is available on desktop and mobile platforms and is available as a free, ad-supported version, as well as a paid subscription version that offers additional features and benefits such as offline listening and higher quality audio. Spotify is available in countries worldwide and has a large user base, making it one of the most popular music streaming services on the market [1].

2.2Anghami

Anghami is a music streaming service that is available on mobile devices in the Middle East and North Africa. It offers a large library of music and allows users to Generate and share music lists, uncover fresh tracks through tailored suggestions, and tune in to online radio stations. Anghami is available as a free, ad-supported version, as well as a paid subscription version that offers additional features and benefit. [2].

2.3Deezer:

Deezer is a French-based platform for streaming music online. It enables users to access music tracks from multiple record labels and podcast programs on their chosen devices, both online and offline, the more songs you listen to, the more the Flow algorithm will learn what music you like. So, the only thing for it is to stream more music. Listen to a variety of different favorite artists and genres and the recommendations you're given should get better as a result [3].

2.4Listen to:

listenTo is platform for songs and music provided with AI that recognize the songs that you are interested by using short videos created to know the user's taste better and let the user use this platform most of the time and also the artist can post on his profile and users can comment and like the post or he can put party tickets in the platform.

Table 2.1: Major features in relevant Web site

Features	Spotify	Anghami	Deezer	Listen to
Music library	Large	Large	Large	Large
Local files	Yes	No	Yes	No
Free version	Yes (ad-supported)	Yes (ad-supported)	Yes (ad-supported)	Yes
Paid version	Yes	Yes	Yes	No
Platforms	Desktop and mobile	Mobile only	Desktop and mobile	Web Site
Available countries	Worldwide	Middle East and North Africa	Worldwide	Worldwide
Podcasts	Yes	No	Yes	No
Lyrics	Yes	Yes	Yes	No
Custom playlists	Yes	Yes	Yes	Yes
Radio	Yes	Yes	Yes	Yes
Social features	Yes	No	Yes	Yes
Artist timeline	No	Yes	No	Yes
Artist creates party	No	No	No	Yes

Chapter 3: System Technologies

In this chapter we will desiccation the Web framework, technologies, databases and Machine learning models that will be used in build the system.

Web development is the process of building and maintaining websites and web applications. It involves a wide range of activities, including design, programming, testing, and deployment.

3.1 Framework and programming Language:

Web developers use a variety of technologies and frameworks to create dynamic and interactive web experiences. These technologies include HTML (HyperText Markup Language) for structure and formatting, CSS (Cascading Style Sheets) for styling and layout, and JavaScript for interactivity and functionality. Other common technologies used in web development include server-side languages such as PHP and database management systems such as MySQL.

Web developers may work on the front-end (client-side) or back-end (server-side) of a website or application. Front-end developers are responsible for the design and layout of a website, as well as the user experience. They may use frameworks such as React or Angular to build interactive and responsive interfaces. Back-end developers focus on the server-side of web development, building and maintaining the infrastructure and logic that powers a website or application.

The field of web development is always changing with the introduction of new technologies and advancements in best practices. In order to produce exceptional and functional websites and applications, web developers must keep abreast of the most recent tools and techniques. Keeping informed is imperative for success in web development.

Overall, web development is an essential part of the modern digital landscape, and is involved in the creation of websites and applications that are used by millions of people every day. Whether you are building a simple personal website or a complex web application, the skills and knowledge of a web developer are critical to the success of the project.

For web application native language:

- HTML
- CSS
- JS
- PHP

For web Framework used in Backend :

- Java SpringBoot : Spring Boot is a popular Java-based framework used in web development to create stand-alone, production-grade applications. It provides a number of useful features and tools to make it easier and faster to develop web applications. Spring Boot uses dependency injection to manage the components of an application, and includes a built-in server, so applications can be deployed and run without the need for additional infrastructure. It also offers auto-configuration, which can help to reduce the amount of configuration required to set up a new application. Overall, Spring Boot is a powerful and convenient tool for developing web applications in Java [4].

- C# ASP.Net: ASP.NET is a web development framework developed by Microsoft and written in C#. It is used to build dynamic web applications and is a popular choice for developers who work with the .NET ecosystem. ASP.NET provides a number of tools and features to help developers build scalable and secure web applications, including a rich set of controls and libraries, a powerful runtime environment, and integration with other Microsoft technologies. It also supports a variety of programming models, including MVC (Model-View-Controller), Web Forms, and Razor pages, allowing developers to choose the approach that best fits their needs [5].

- PHP Laravel: Laravel is an open-source PHP framework used for web application development. It is known for its elegant syntax and tools for tasks such as routing, authentication, and object-relational mapping. Laravel also includes a number of features to help developers build scalable and maintainable applications, such as a built-in task scheduler, a powerful queue system, and support for unit testing. Laravel is a popular choice for developers due to its ease of use, extensive documentation, and active community of contributors. It is often used to build web applications that need to handle large amounts of data or require complex business logic [6].

- Python djanko: Django is a high-level Python web framework used for the rapid development of secure and maintainable web applications. It follows the model-template-view architectural pattern and provides a number of features to help developers build scalable and efficient web applications, including a powerful ORM (Object-Relational Mapper) for interacting with databases, a built-in template engine, and support for asynchronous request handling. Django is known for its emphasis on security and reliability, and is often used to build complex, data-driven web applications. It is a popular choice for developers who prefer to work with Python [7].

Spring Boot, ASP.NET, Laravel, and Django are all popular frameworks used for web application development, but they have some key differences:

- Language: Spring Boot is written in Java, while ASP.NET is written in C#, Laravel is written in PHP, and Django is written in Python. This means that developers who are familiar with one of these languages will likely find it easier to work with the corresponding framework.

- Architecture: Spring Boot is designed to be a standalone framework that can be used to build a wide range of applications. It uses dependency injection to manage the components of an application and includes a built-in server, so applications can be deployed and run without the need for additional infrastructure. ASP.NET is built on top of the .NET framework, which provides a runtime environment and a set of libraries for building applications. Laravel is a PHP framework that is designed for building web applications, but it can also be used for other types of projects. Django follows the model-template-view architectural pattern and is designed for the rapid development of secure and maintainable web applications.

- Features: Spring Boot offers a number of useful features to help developers build and deploy applications more efficiently. These include auto-configuration, which can help to reduce the amount of configuration required to set up a new application, and a built-in server, which eliminates the need to set up and configure a separate web server. ASP.NET provides a rich set of controls and libraries, a powerful runtime environment, and integration with other Microsoft technologies. It also supports a variety of programming models, including MVC, Web Forms, and Razor pages, allowing developers to choose the approach that best fits their needs. Laravel includes tools for tasks such as routing, authentication, and object-relational mapping, as well as features such as a built-in task scheduler and support for unit testing. Django provides a powerful ORM (Object-Relational Mapper) for interacting with databases, a built-in template engine, and support for asynchronous request handling. It is known for its emphasis on security and reliability.

Overall, these frameworks offer different approaches and features for web application development, and the best choice will depend on the needs of the project and the preferences of the developer. Spring Boot is a convenient choice for developers who want to build standalone applications quickly and easily, while ASP.NET is a powerful framework that is well-suited to enterprise-level applications. Laravel is a popular choice for developers who want to build web applications using PHP, and offers a number of useful tools and features to help them do so. The high-level Python web framework Django is perfect for creating secure, maintainable web applications quickly.

Table 3.1: Springboot vs ASP.NET Vs Laravel vs Django [14-17]

Features	Spring Boot	ASP.NET	Laravel	Django
Language	Java	C#	PHP	Python
Standalone	Yes	No	No	No
Dependency	Yes	No	Yes	No
Built-in server	Yes	No	No	No
Auto-config	Yes	No	No	No
Programming	N/A	MVC, Web	N/A	MVC
Forms, Razor		Forms, razor pages		
Task scheduler	No	No	Yes	No
Queue system	No	No	Yes	No
Unit testing	No	No	Yes	Yes
ORM	Yes	No	No	Yes
Template engine	No	No	No	Yes
Async upport	No	No	No	Yes

Ultimately, our decision was to implement Django as our web framework. This is due to its rapidly growing popularity globally and its powerful yet intuitive design, as well as its comprehensive documentation. Additionally, Django is built on Python, which is similar to the programming languages we are already familiar with.

3.2 Databases:

There are two ways of store data in databases

1- SQL (Structured Query Language) a

2- NoSQL (Not Only SQL)

They are types of database management systems. SQL databases are based on a traditional relational model and use structured query language for defining and manipulating the data. They are well-suited to handling large amounts of structured data and are commonly used in applications that require complex queries and transactions. NoSQL databases, on the other hand, are designed to handle large amounts of unstructured data and can scale horizontally across multiple servers. They often use a variety of data models, such as key-value, document, columnar, and graph, and are commonly used in applications that require high performance, large data volume, and real-time processing.

We will use MySQL databases management in our system.

MySQL is a popular open-source database management system that is often used in web applications to store and manage data. It is based on the SQL language and is known for its reliability, performance, and ease of use. MySQL can be used with a variety of programming languages and web frameworks, making it a convenient choice for developers building web applications.

In a web application, MySQL can be used to store a wide variety of data, such as user profiles, product information, and blog posts. It supports a range of data types, including numbers, strings, and binary data, and offers features such as transactions, stored procedures, and triggers to help developers build robust and scalable applications. MySQL can be accessed via a number of languages and libraries, including PHP, Python, and Ruby, making it easy to integrate with a variety of web development frameworks.

MySQL is commonly used in combination with server-side languages such, python and Ruby on Rails, and is well-suited to handling large amounts of structured data. It can scale to support large numbers of users and transactions, and is often used in applications that require complex queries and transactions. Overall, MySQL is a versatile and reliable choice for storing and managing data in web applications [8].

3.3 System Architecture and State Management

The architecture of a web application is a blueprint that showcases the relationship between the components, middleware systems, user interfaces, and databases. This interconnection enables multiple applications to function simultaneously.

Web application architecture outlines the simultaneous interactions between components, servers, user interfaces, middleware systems, and databases in a web application. It serves as a structural layout, defining the logical connection between the client and server sides for an optimal web experience.

When a user accesses a webpage, the server sends specific data in response to the user's request. More specifically, a web client, also known as a user agent, can request web resources or documents, such as HTML, JSON, or PDF, from a web server. Once these minimal actions are completed, the requested information is displayed. This marks the beginning of the interaction between the user and the website.

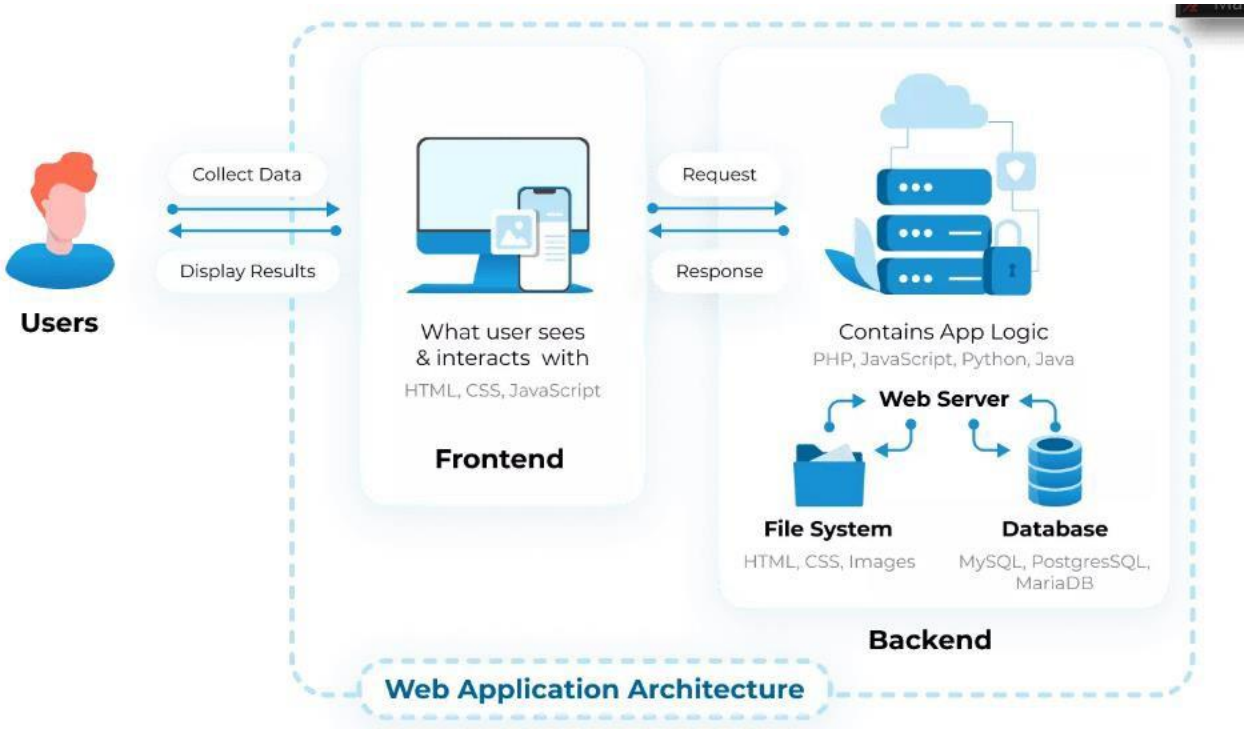


Figure 1: Web Application Architecture

3.4 Version Control System and Repository

We used Git, so we can go back to any version of our project, and GitHub so we make it easier for the members to contribute together on the same project without needing to hard copy each line of code to each project, so it's like we created one project and we pushed it online than we contributed to it together.

3.5 Task Management

We utilized Trello, a visual tool that enables teams to manage projects, workflows, and tasks by adding files, checklists, and automation. We used Trello to deploy Scrum, an agile software development technique that makes use of incremental and iterative procedures. Scrum is a highly flexible, quick, and successful structure that offers clients value throughout the course of the project.

Chapter 4: Software Architecture

System Requirement Specification (SRS)

In this section we will present a detailed description about listenTo System, it will have details about what the system shall do and how it is expected to do it.

The listenTo System will have two parts: a web-based portal for basic tasks of the Admin, Artist and normal users' management system, and this part will be used by the end users. The other part in the back end will have an AI algorithm that will help the system to suggest best songs for the user and close to their taste.

4.1 Overall Description:

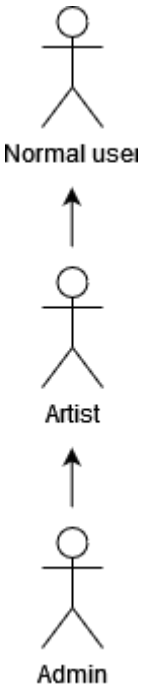
The listenTo System will have the following basic functionalities:

1. User creation and user roles.
2. Artist can create an Artiest role.
3. Admin accepts artiest.
4. Artist can add songs.
5. Artist can post on his timeline.
6. Artist can create an album.
7. User can comment on an artist timeline.
8. User can listen to the artist's songs.
9. User can rate a song.
10. User can rate an artist.
11. User can search about the songs.
12. User can search about artists.
13. User can create a playlist.
14. The system can recognize a song.

4.2 Group Specification:

The listenTo System will be used mainly by the following user groups:

1. The system administrators (owners)
2. The Artist members.
3. The normal members.



4.3 User Requirements:

User requirement describes users' needs from the system; it is written before the system is created. Since it provides a better sense of how the system will operate. User requirements in our software are:

1- The system shall provide a functionality to allow registering into the website
2- The system shall allow listening to the songs
3- The system shall allow the artist to upload songs
4- The system shall allow the user to create album
5- The system shall provide a functionality to search for an artist
6- The system shall provide a functionality to search for a song
7- The system shall provide an admin panel
8- The system shall learn about the user taste
9- The system shall suggest songs for the user
10- The system shall provide a functionality to rate
11- The system shall provide songs recognition

Table 4.3.1: Normal members group description

Group Name	Normal Members
Description	The people who have normal account to interact with songs and Book party tickets that posted by artists, rating the artist based on his experience, like or dislike a song.
Profile Data	first name, last name, birth date, gender, email, mobile, address.
Super-group	Admin
Relevant use cases	Login, search, view artist profile, rating artist, interact with songs

Table 4.3.2: Artist members group description

Group Name	Artist members
Description	The people who have Artist account that accepted by the admin, artist create a party and selling a ticket, crate an album and post songs on it And see the user's interaction of songs
Profile Data	first name, last name address , gender email , mobile , birth date , songs type.
Super-group	Admin
Relevant use cases	View other artist profile, songs and interact with them, enter party details, provide albums' with songs, view interact with songs history

Table 4.3.3: Administrator group description

Group Name	Administrator
Description	Accent artist account, manage a parties, financial management, view and edit all accounts in the system and control all the system
Profile Data	First name, last name, email, mobile.
Super-group	Non
Relevant use cases	Search members, add member, add artist account , manage events in the system and stages ,

4.4 System Requirements:

The listenTo System provide some of the basic functions for the user groups, the following is the details of the functional requirements for groups.

1 UR1: The system members shall be able to login to the system.
- 1.1 The system shall provide a login form to allow the member to login using his email and password.
- 1.2 The system shall authenticate the users using the existing information of members and a password.

2 UR2: The normal member shall be able to search for an artist and songs provider.

- 2.1 The system shall display a list of artist services providers.
- 2.2 The system shall provide search filters to search for artist services providers.
- 2.3 The member shall enter at least one search criterion.
- 2.4 The search results shall be refined using the search criteria the member entered.
- 2.5 The search results shall include the artist service name, the services types, and the contact information

3 UR3: The artist member shall be able to see the rates and feedback details.

- 3.1 The system shall provide a link for the member to view the rate for each song
- 3.2 The system shall provide a link for the member to view the rate for each event
- 3.3 The system shall provide a link for the member to view the rate users it self

4 UR4: the user can register to the system

- 4.1 the system shall allow the user to enter his details like e-mail, password, full name and phone number
- 4.2 the system shall create a user profile from the available data
- 4.3 the system shall save the data in the database

5 UR5: The artist member can upload or delete a song

- 5.1 the system shall allow the artist to upload songs with description.
- 5.2 the system shall allow the artist to write song name
- 5.3 the system shall allow the artist to write song description
- 5.4 the system shall allow the artist to write song type
- 5.5 the system shall allow the artist to write song lyrics
- 5.6 The system shall allow the artist delete a song.

6 UR6: the artist shall be able to post and announce a party on his profile
- 6.1 The system shall the system shall allow the user to post a post.
- 6.2 The system shall allow the user to post a calendar to its parties.
- 6.3 The system shall allow the user to edit and delete his posts and parties.

7 UR7: the system shall allow the artist to create an album .
- 7.1 The system shall allow the artist to create album name.
- 7.2 The system shall allow the artist to add songs to the album

8 UR8:The system shall provide a functionality to search for an artist

- 8.1 The system shall allow the artist to search on a song by writing its name.
- 8.2 The system shall allow the artist to search about other artist or users by their names.
- 8.3 The system shall allow the artist to search about albums.
- 8.4 The system shall allow the artist to search about public playlists.

9 UR9: an artist can follow other artist or users.
- 9.1 The system shall allow the artist to follow other artist by clicking follow on artist's profile.
- 9.2 The system shall allow the artist to follow other users by clicking follow on user's profile.

10 UR10: The administrator shall have control overall the website.

- 10.1 The system shall allow the admin to delete any user.
- 10.2 The system shall allow the admin to accept any artist request.
- 10.3 The system shall allow the admin to delete any song.
- 10.4 The system shall allow the admin to delete any event.
- 10.5 The system shall allow the admin to reject any artist request.
- 10.6 The admin can delete any public play list.
- 10.7 The admin can delete any album.
- 10.8 The admin can suspend a user.
- 10.9 The admin can suspend an artist.

11 UR11: The administrator shall be able to search for artist and users and songs.

- 11.1 The system shall allow the admin to search for an artist
- 11.2 The admin can search for a user.
- 11.3 The system shall allow the admin to search for public play list.
- 11.4 The system shall allow the admin to search for an album.

12 UR12: the system shall provide a suggestion system.

- 12.1 The system shall provide a special suggestions for each user depends on his taste of music.

4.5. Use Cases:

The following is the system overall use case diagram:

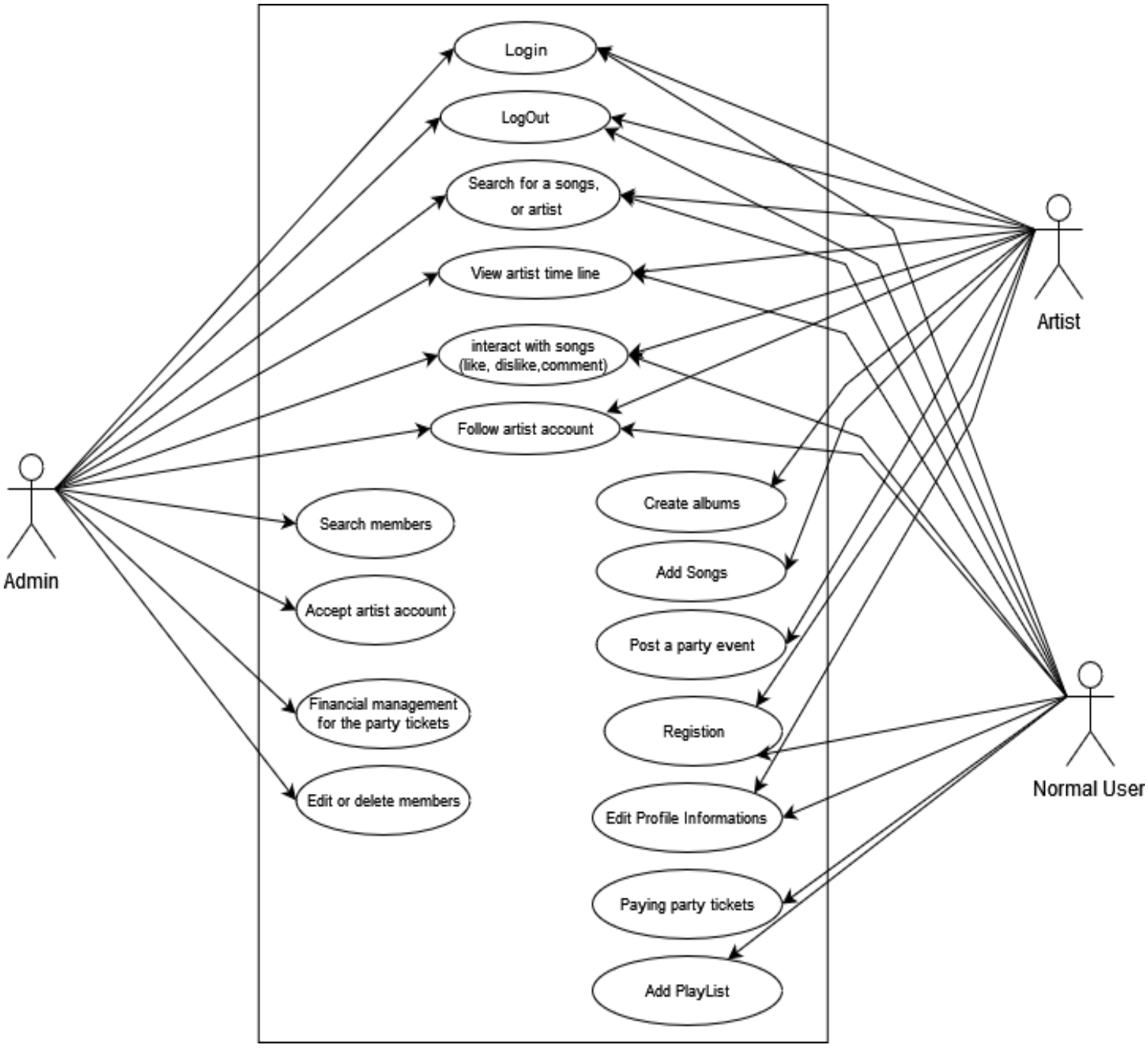


Figure 3: Use Case Diagram

4.6 Class Diagram:

The following diagram shows the initial class diagram for the system, it shows the basic attributes that we need, and the operations that will be performed will be based on these attributes and will be determined later when we start the application programming.

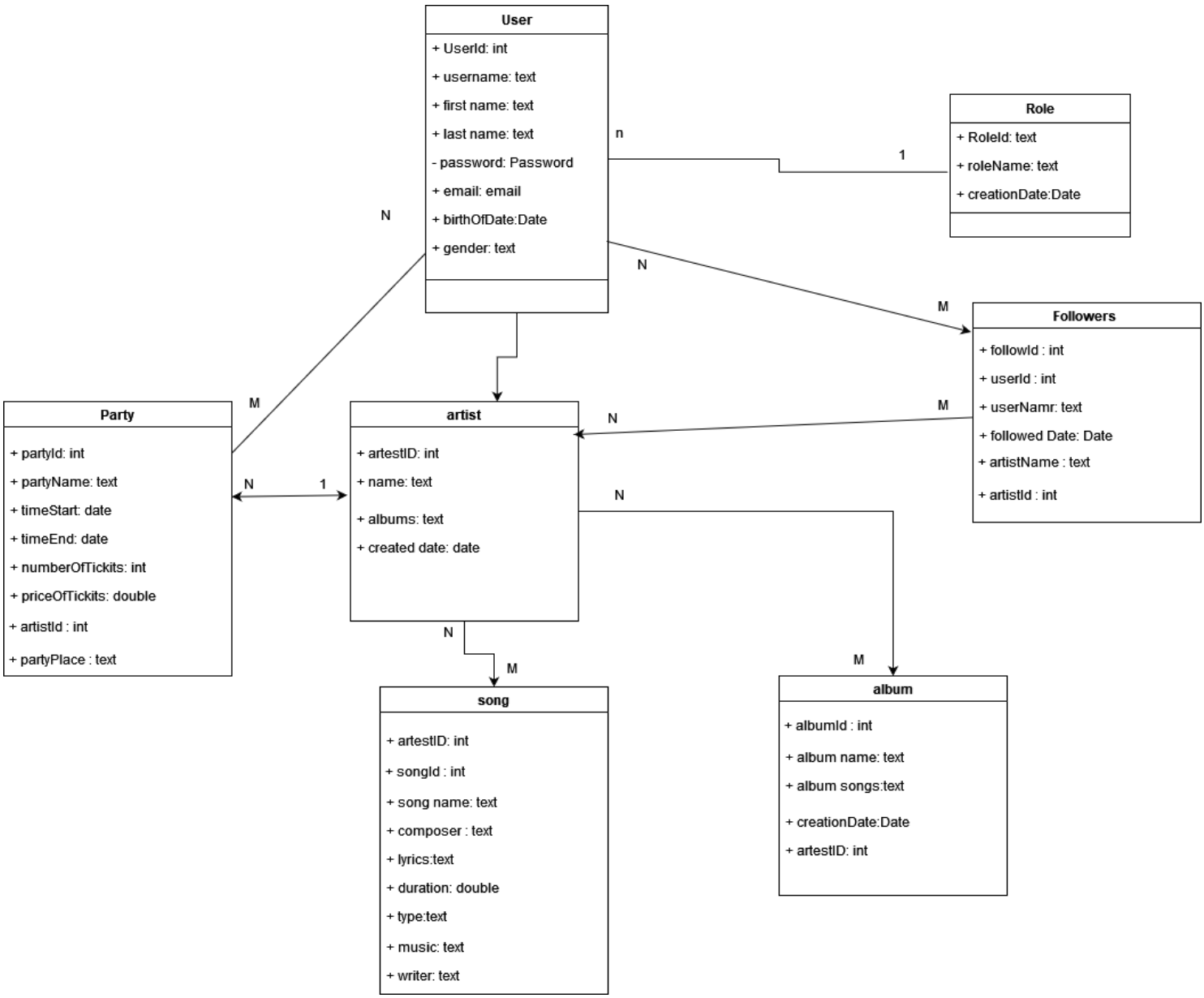


Figure 4: Class Diagram

4.7 Activity Diagram:

An activity diagram is a type of behavioral diagram that illustrates the behavior of a system. It displays the control flow from the starting point to the end point, displaying the different decision paths that exist during the execution of the activity. [9].

4.7.1 **Signup and Login :**Every app begins with registration (sign up) and login. While login is the process of authenticating registered users to their accounts, signup is the process of creating an account for users in the system. Figure 4. shows sign up and login processes.

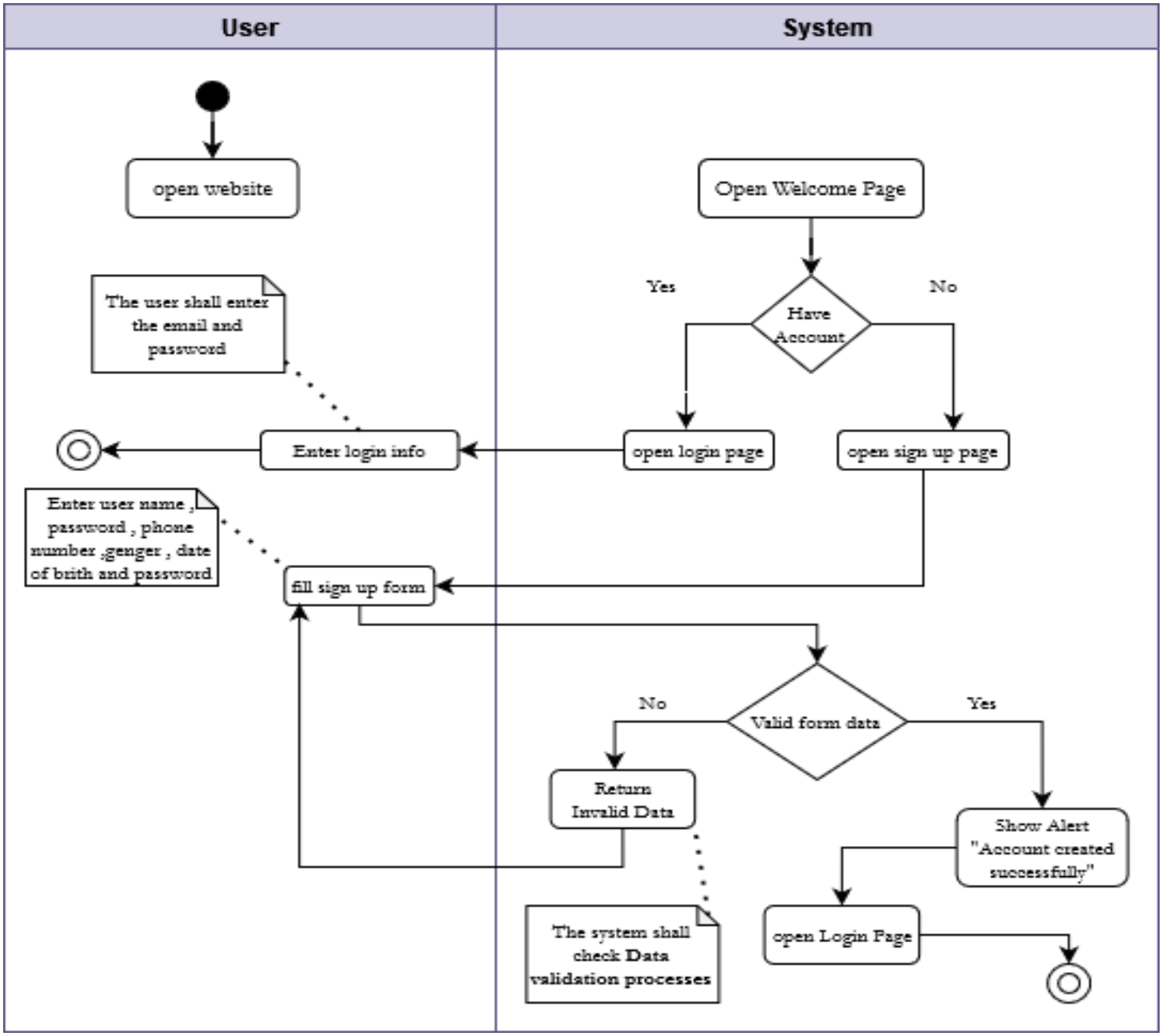


Figure 4: Signup and Login Activity Diagram

4.7.2 **Create artist account**

A sign in and registration of the artist and song upload the artist can sign up and sign in like the normal user because they will sign up like a normal user then they will ask to be an artist and the adman can approve or decline the application, so the artist can reserve for an event.

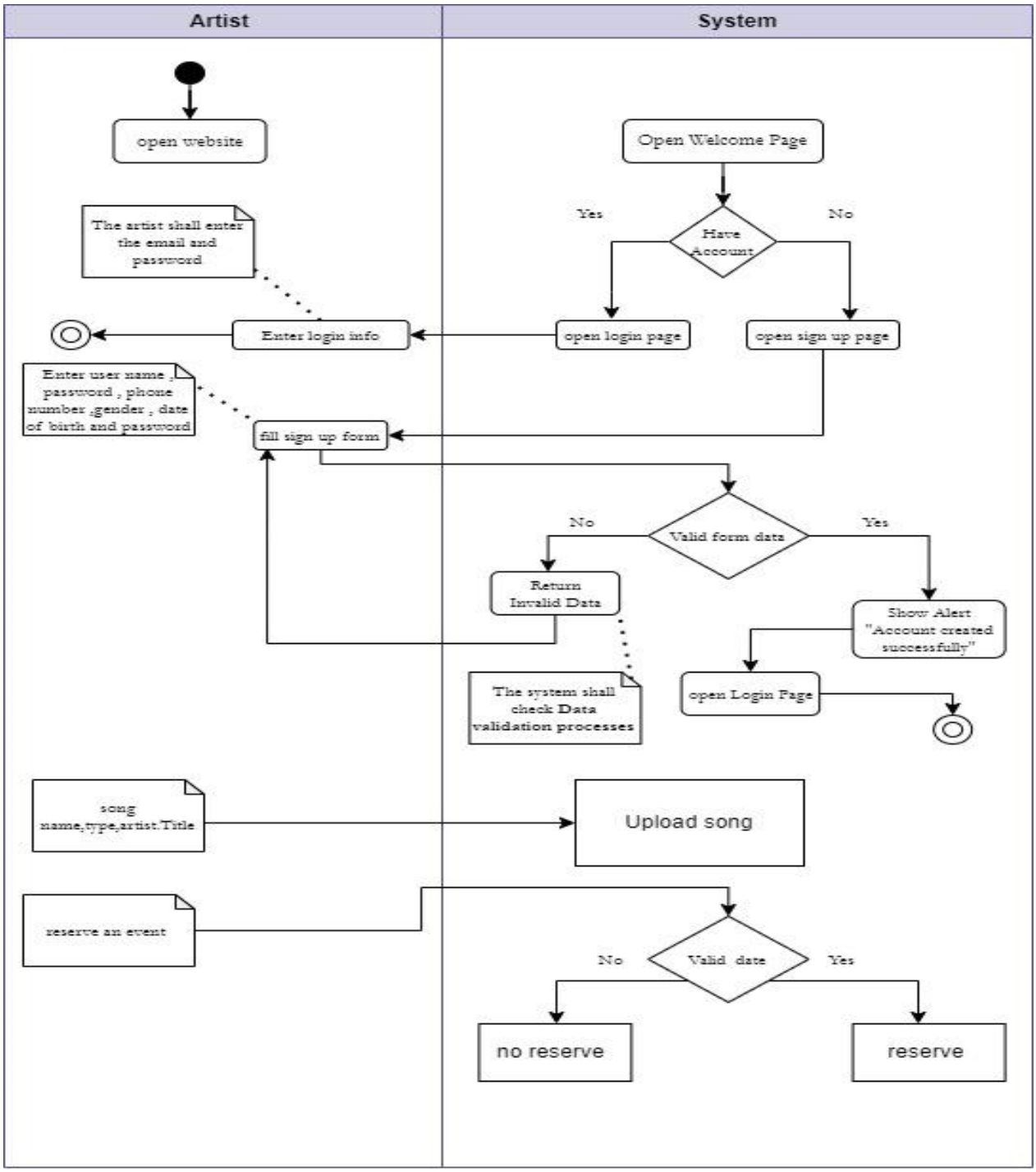


Figure 5: Create artist account Activity Diagram

4.7.3 Search for a song, artist and event

Searching in web site is a core feature in that site. The users will be able to search for any features effortlessly and filtering the results by using keys such as artist name or songs time and so on. Figure 4.4 shows searching doctor process

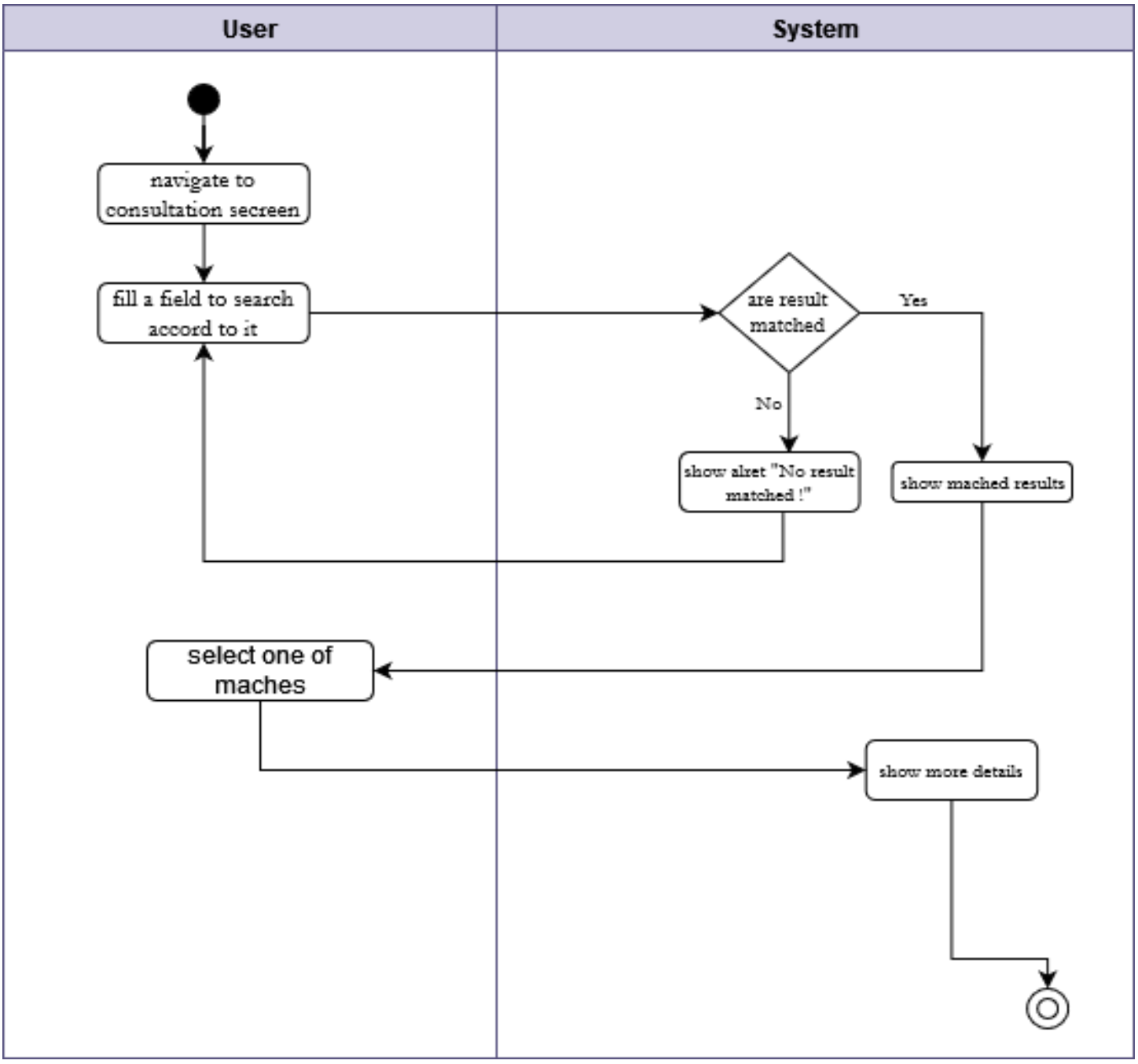


Figure 6: Search for a song, artist and event Activity Diagram

4.7.4 Artist booking an event

Artist Booking an event:

The artist shall book an event and the system will check if there's any slots available if the system didn't found any slots will return a message for the artist else the system going to book an event and display it to the users and the system automatically will cut the profit to 20% (for the system) and 80% (for the artist)

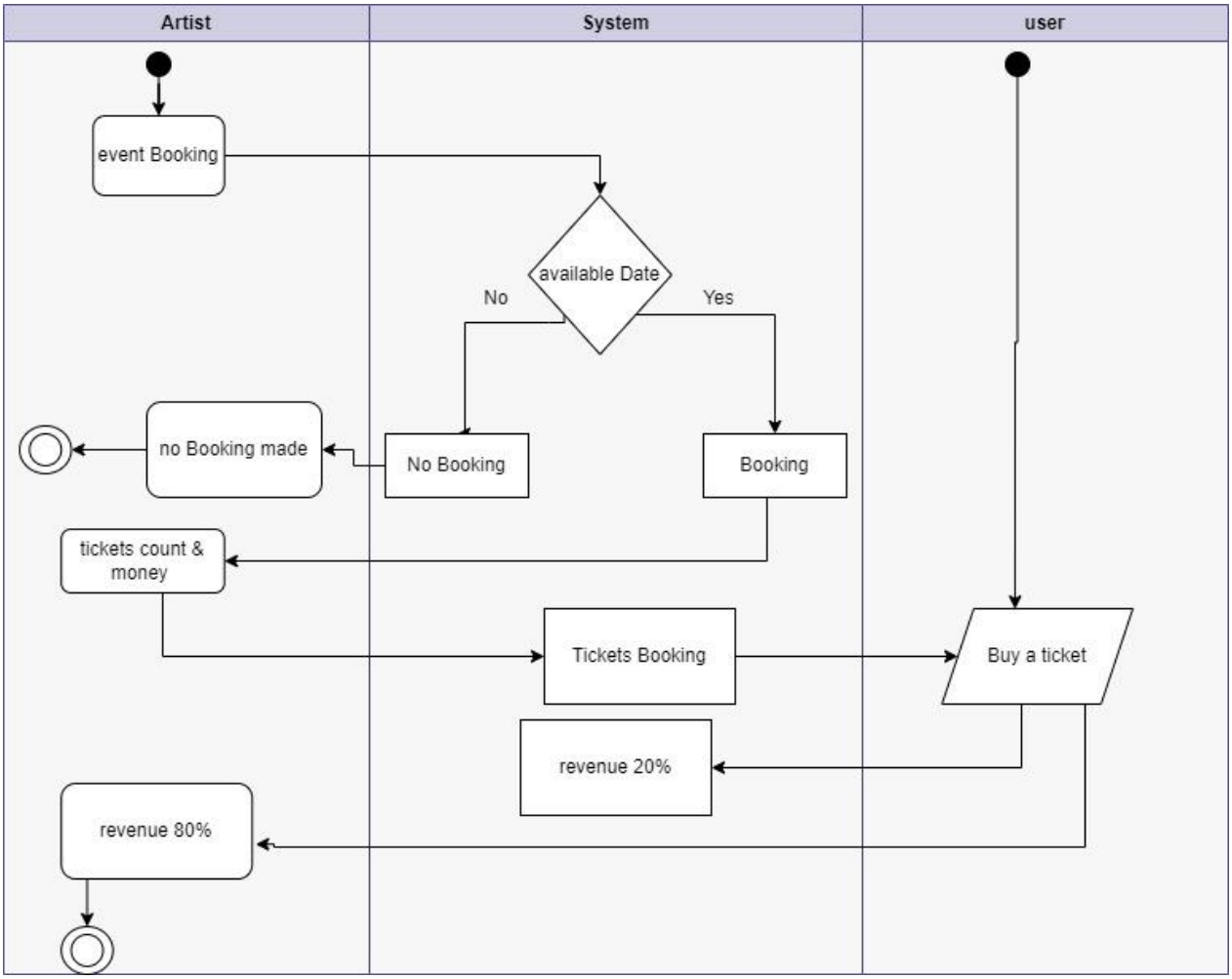


Figure 7: Artist booking an event Activity Diagram

4.8 Machine Learning Model and Recommender Systems

K-means clustering is a prevalent unsupervised machine learning technique that identifies patterns in data by dividing it into a pre-determined number of clusters, each represented by its mean (centroid). The algorithm operates by iteratively assigning each data point to the closest centroid, utilizing a distance metric such as Euclidean distance. After all data points have been assigned, the centroids are recalculated as the average of the points in their respective clusters. This iteration continues until the centroids no longer change, indicating convergence.

One of the main advantages of k-means clustering is that it is relatively fast and easy to implement, especially for large datasets. It is also highly scalable, meaning that it can handle very large datasets with ease. Additionally, it is relatively easy to interpret the results of k-means clustering, as it produces clear and concise clusters that can be easily visualized and analyzed [10].

However, k-means clustering also has some limitations. One major limitation is that it requires the number of clusters to be specified in advance, which can be difficult to determine in some cases. Additionally, the algorithm is sensitive to the initial placement of the centroids, and can produce different results depending on the initialization method used. For example, if the centroids are initialized in a poor location, the algorithm may converge to a suboptimal solution. Finally, k-means clustering is sensitive to outliers and may produce poor results if the data contains a large number of outliers.

Despite these limitations, k-means clustering remains a popular and useful tool for a wide range of applications, including recommendation systems. It is particularly useful for identifying patterns and trends in large datasets, and can be an effective way to create customized suggestions tailored to individual users based on their prior actions and inclinations. In a recommendation system, k-means clustering can be used to group users into different segments, and then generate recommendations based on the products that are popular among users in the same cluster. This can help to ensure that the recommendations are more tailored to the individual user, and therefore more likely to be relevant and useful [11].

4.8.1 Data Set and feature extraction

Dataset for Music Recommendation System:

The model's training dataset was sourced from a collection of 10,000 popular and genre-diverse mp3 songs, in order to ensure that the data was representative and diverse. This dataset was collected and preprocessed to extract various features that describe each song. The feature extraction process is a crucial step in building a recommendation system, as it determines the quality and reliability of the final model. In this report, we will discuss the feature extraction process and the features obtained from it, as well as the data analysis that was performed to choose the k-means algorithm for building the recommendation system.

Feature Extraction:

Feature extraction is the process of extracting relevant information from raw data and transforming it into a set of features that can be used to train a machine learning model. In the case of music recommendation systems, the raw data is the MP3 songs, and the features are the characteristics of each song that can be used to compare and recommend songs.

One commonly used feature extraction method in Python is the librosa library, which provides a wide range of audio analysis tools. In our case, we used librosa to extract the following features from each MP3 song. (which take 8 days of running 8 process to extract the features from our mp3 set):

These features were selected by the algorithm used in python, then Upon processing the data, conducting a comprehensive analysis, and engaging in rigorous evaluation, we were able to deduce and carefully select the most relevant characteristics that will be utilized in the proposal process, while also ensuring that any necessary modifications were implemented. These modifications include,

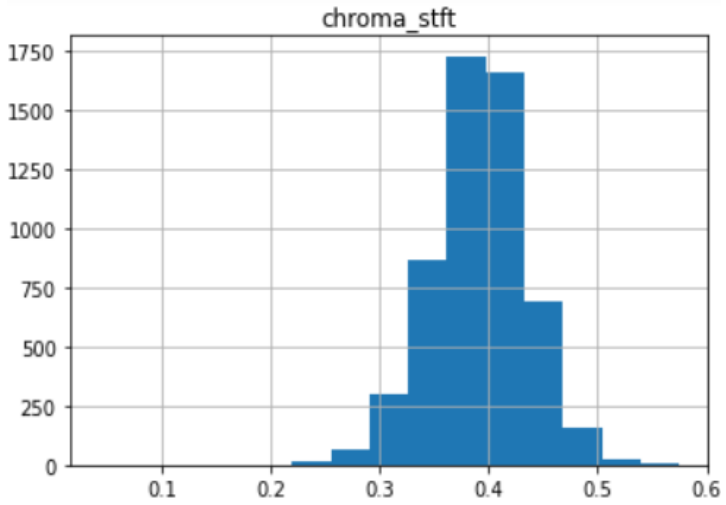
- chroma_stft: used in audio signal processing to represent the pitch content of a sound signal. This technique involves calculating the magnitude of the short-time Fourier transform (STFT) of the signal, followed by mapping the resulting spectrum onto a chromatic scale.
- spectral_centroid: used in audio signal processing to represent the center of gravity of the frequency spectrum of a sound signal. This technique involves calculating the weighted average of the frequency components, with the weighting factor determined by the spectral amplitude
- spectral_bandwidth : used in audio signal processing to represent the width of the frequency spectrum of a sound signal. This technique involves calculating the average distance of the frequency components from the spectral centroid, which provides a measure of the spread of the spectrum
- spectral_rolloff: used in audio signal processing to represent the frequency content of a sound signal. This technique involves calculating the frequency below which a certain

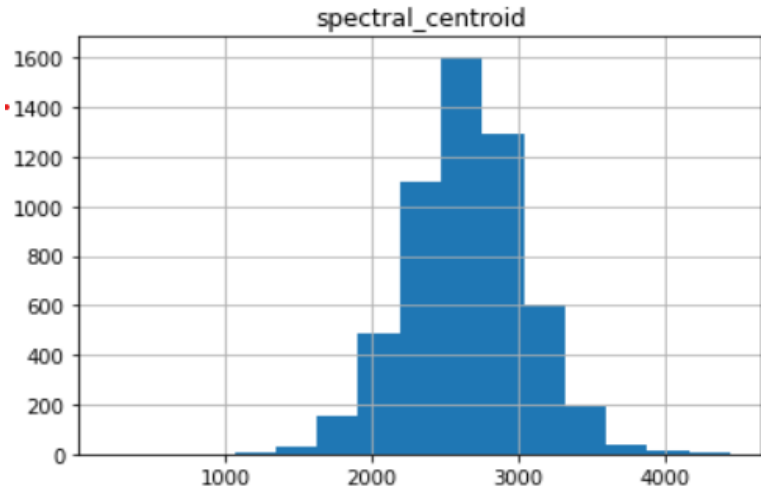
percentage of the total spectral energy is contained, which provides a measure of the spectral shape.

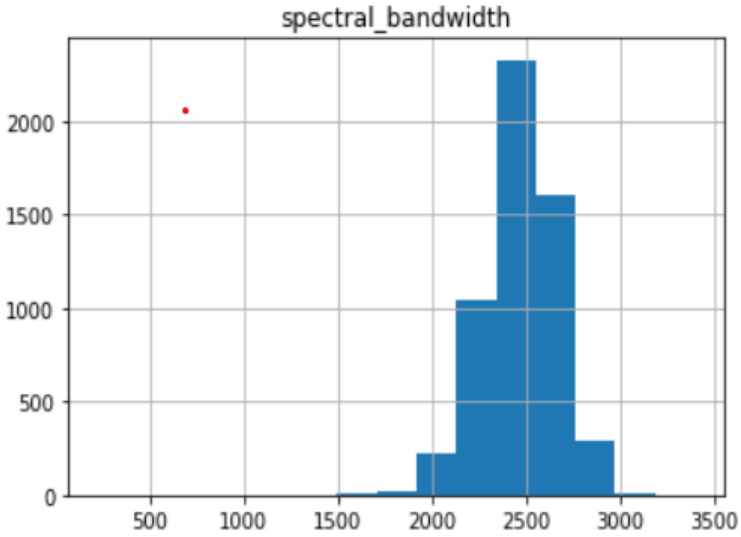
- spectral_contrast: used in audio signal processing to represent the spectral shape of a sound signal. This technique involves calculating the difference in energy between adjacent frequency bands, which provides a measure of the spectral contrast.
- tempogram: used in music signal processing to represent the rhythmic content of a musical piece. This technique involves calculating the autocorrelation of the input signal's onset strength envelope, which provides a measure of the temporal regularity of the rhythm.
- chroma_cqt: used in music signal processing to represent the tonal content of a musical piece. This technique involves calculating the chroma vector of each frame of the input signal using a Constant-Q Transform, which provides a more accurate representation of the frequency content of musical signals. Chroma_cqt features are commonly used in music analysis for tasks such as chord recognition, key detection, and genre classification.
- chroma_cens: used in music signal processing to represent the tonal content of a musical piece. This technique involves calculating the chroma vector of each frame of the input signal, which represents the distribution of musical pitches over the 12-tone chromatic scale.
- melspectrogram: used in audio signal processing to represent the power spectrum of a sound signal on a mel frequency scale. This technique involves dividing the input signal into small frames, computing the power spectrum of each frame using a Fast Fourier Transform, and mapping the resulting frequencies onto the mel scale.
- mfcc: used in speech and audio processing to represent the spectral characteristics of a sound signal. This technique involves calculating the frequency response of the signal and mapping it onto a mel frequency scale, followed by applying a Discrete Cosine Transform to obtain the cepstral coefficients. MFCCs are commonly used for tasks such as speech recognition, speaker identification, and music genre classification.
- poly_features: is a feature extraction technique used in signal processing that involves fitting a polynomial model to the input signal to estimate the relationships between the features. This technique is commonly used in speech and audio processing to identify the characteristics of the input signal and can be used to extract features such as formants and harmonics.
- tonnetz: is a feature used in music signal processing to visualize and analyze musical harmonies by representing pitches as points on a triangular lattice. This feature helps identify relationships between different musical chords and is often used in music theory and composition.
- rms: is a statistical measure commonly used in signal processing to measure the average power of a signal over time. This feature is calculated by taking the square root of the mean of the squared values of the signal over a specified time interval. RMS is often used in audio and vibration analysis to determine the intensity or energy of a signal.
- zero_crossing_rate: Zero crossing rate is a feature used in audio signal processing to estimate the frequency content of a signal by counting the number of times it crosses the horizontal axis or zero line within a given time interval. This feature can also be used to detect the onset of certain sounds and is commonly used in speech recognition and music analysis.

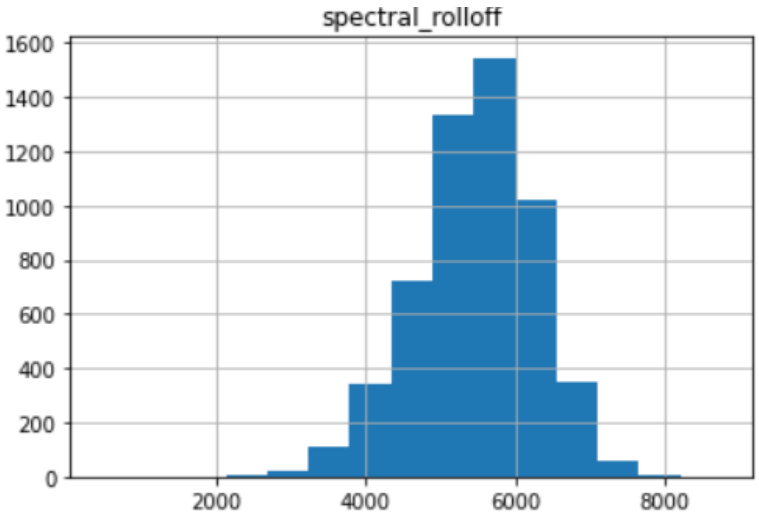
Distribution of features according to values

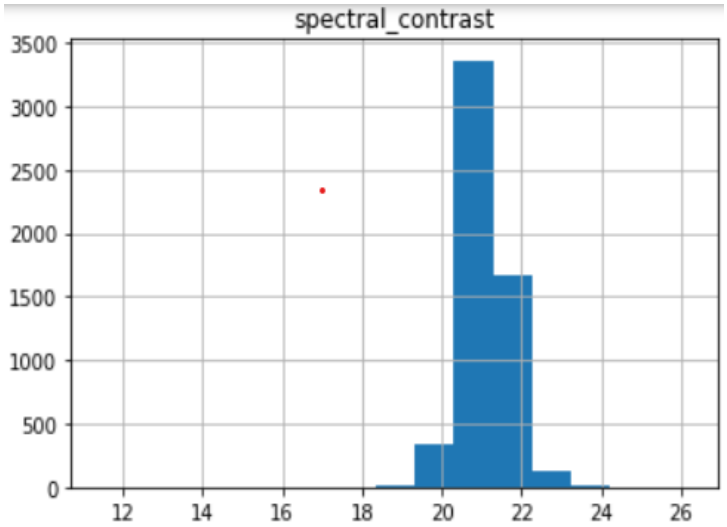
(Visualize the distribution of values in each column of a pandas data frame using histograms.**)**

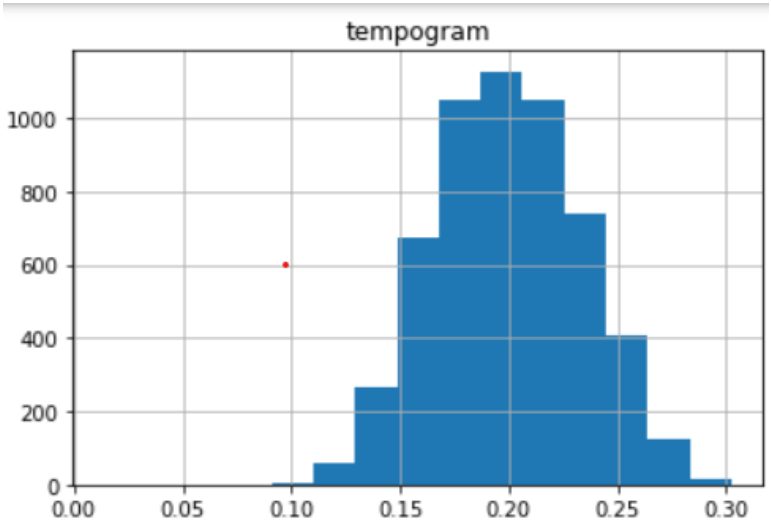


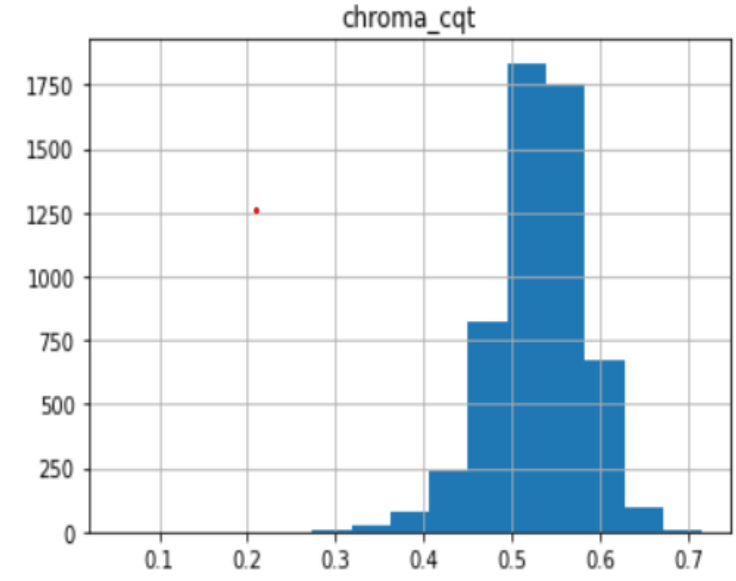


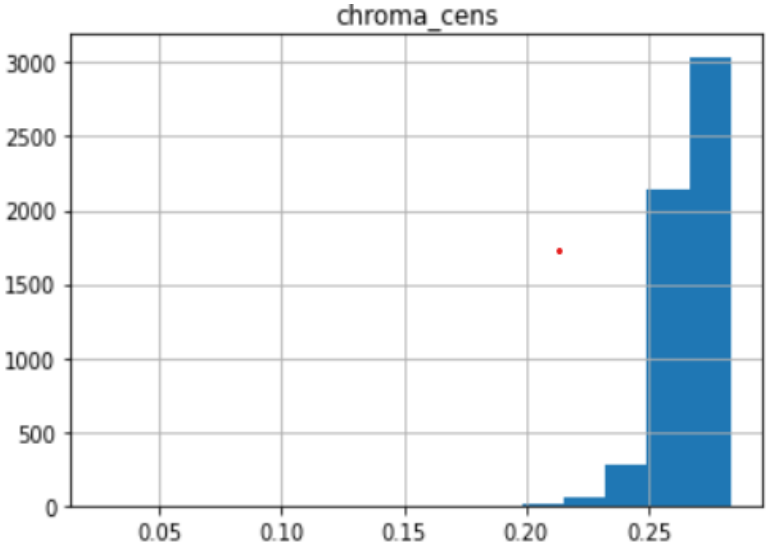


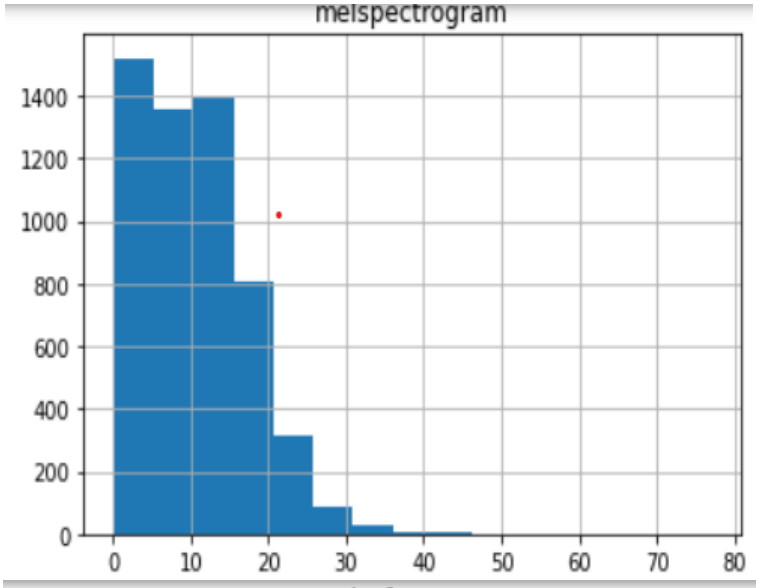


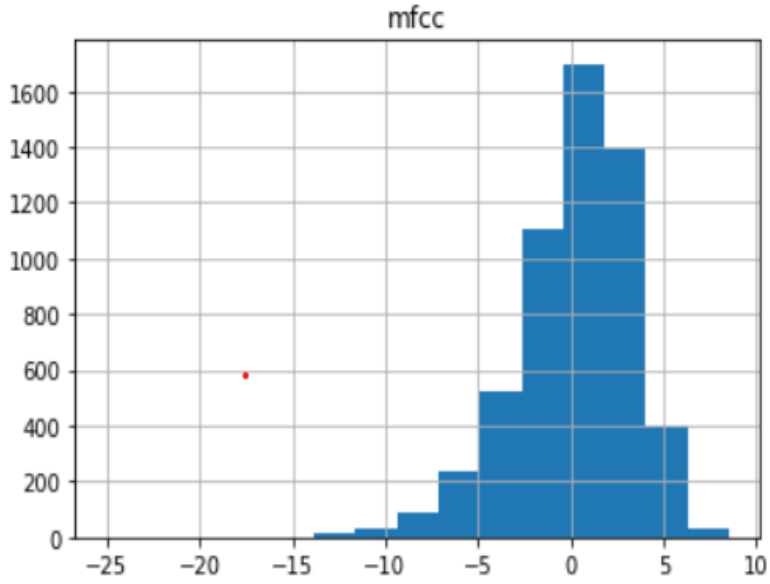


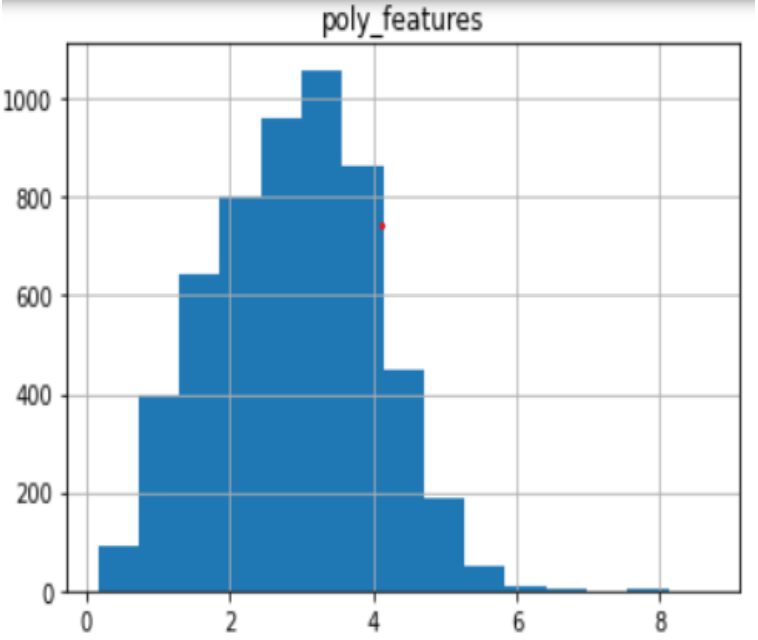


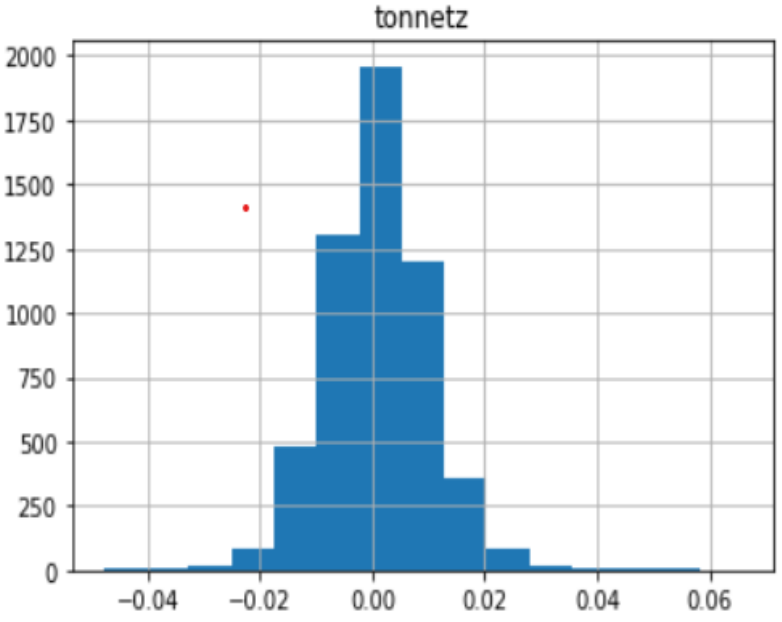

chroma_cqt
1750
1500
1250
1000
750
500
250
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7


chroma_cens
3000
2500
2000
1500
1000
500
0
0.05
0.10
0.15
0.20
0.25


melspectrogram
1400
1200
1000
800
600
400
200
0
0
10
20
30
40
50
60
70
80


mfcc
1600
1400
1200
1000
800
600
400
200
0
−25
−20
−15
−10
−5
0
5
10


poly_features
1000
800
600
400
200
0
0
2
4
6
8


tonnetz
2000
1750
1500
1250
1000
750
500
250
0
−0.04
−0.02
0.00
0.02
0.04
0.06

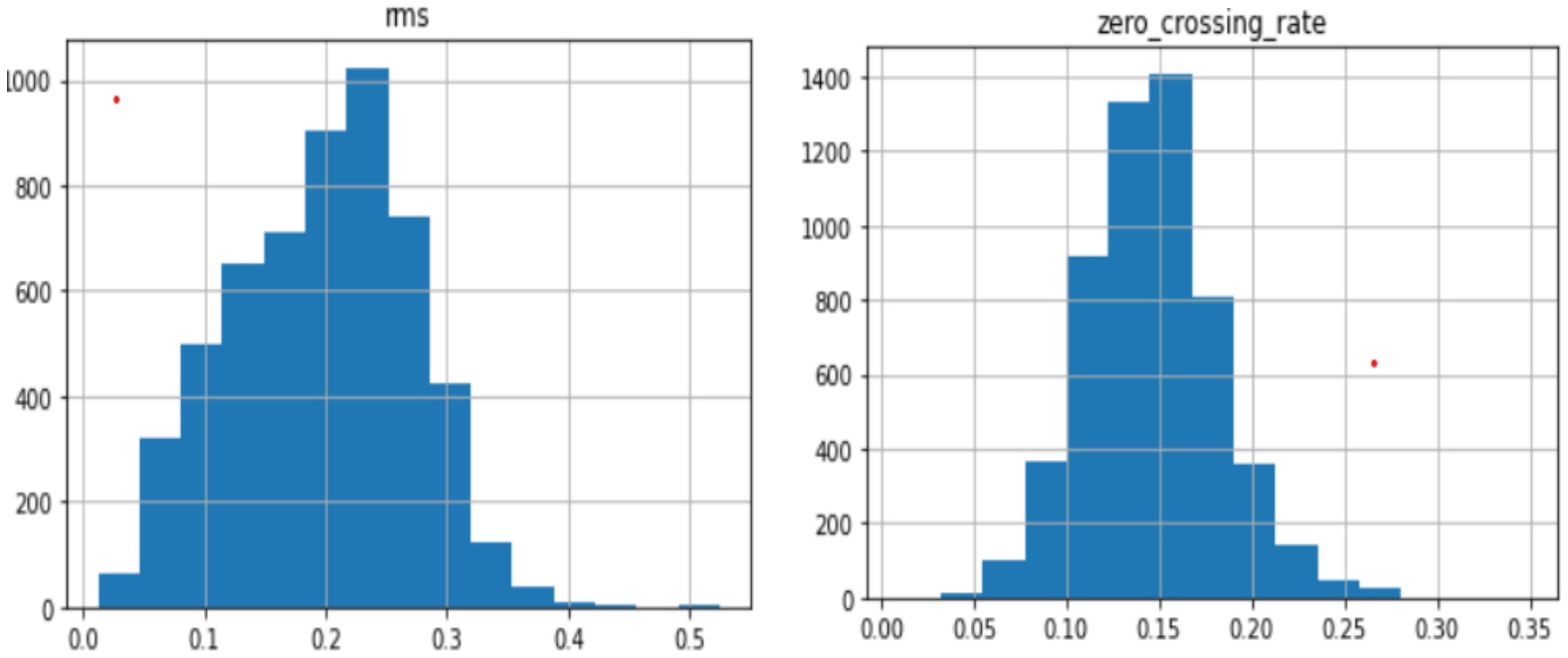


Figure 8: features Data representation

Skewness is a statistical measure of the degree of asymmetry of a probability distribution. In other words, it measures how much the tails of a distribution are stretched out. A normal distribution has zero skewness, meaning its tail is perfectly symmetric. Positive skewness, on the other hand, indicates that the distribution has a long right tail, while negative skewness implies a long left tail.

Skewness can have a significant impact on the training process of a machine learning model. For example, a positively skewed dataset may cause a model to overestimate the effect of high values and underestimate the effect of low values. This is because the majority of the data is concentrated on the left side of the distribution, causing the model to focus more on the high values.

Conversely, a negatively skewed dataset may cause a model to overestimate the effect of low values and underestimate the effect of high values. This is because the majority of the data is concentrated on the right side of the distribution, causing the model to focus more on the low values.

To address this issue, various techniques can be employed to transform the dataset to achieve a more normal distribution. One common technique is to apply a MinMax scaling It scales the data to a fixed range, usually between 0 and 1. This is achieved by subtracting the minimum value from each value in the dataset and then dividing by the range of the dataset. which can help reduce the skewness and make the distribution more symmetric [12].

We try using Square root transformation, logarithmic transformation and MinMax , but in our dataset the best is MinMax because we don’t want to change the features distribution in another side, MinMax solve this case by scaling the without change the distribution

Dataset:

Out[25]:

	chroma_stft	spectral_centroid	spectral_bandwidth	spectral_rolloff	spectral_contrast	tempogram	chroma_cqt	chroma_cens	melspectrogram	mfcc	poly_features	tonnetz	rms	zero_crossing_rate
0	0.601149	0.567253	0.704818	0.599222	0.709194	0.648032	0.727643	0.935546	0.161840	0.768913	0.351149	0.309264	0.411359	0.397000
1	0.553670	0.593203	0.676351	0.590348	0.658004	0.575849	0.724205	0.958179	0.185064	0.727174	0.423799	0.456993	0.487602	0.451821
2	0.603023	0.580563	0.735441	0.644001	0.649170	0.697881	0.670970	0.904126	0.030395	0.635866	0.137519	0.323837	0.169175	0.370741
3	0.708381	0.629202	0.738539	0.646479	0.632431	0.702439	0.763156	0.943765	0.048290	0.740973	0.168955	0.424954	0.202896	0.435395
4	0.702525	0.490909	0.670732	0.532629	0.672041	0.682115	0.822314	0.958503	0.105907	0.775030	0.280997	0.410895	0.313903	0.350490
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
5513	0.551321	0.575064	0.677518	0.592814	0.642961	0.512316	0.681140	0.923903	0.143219	0.722300	0.345089	0.309185	0.403741	0.394760
5514	0.774479	0.722353	0.780831	0.733055	0.595923	0.652896	0.849788	0.970427	0.107918	0.823712	0.356395	0.414689	0.373136	0.611671
5515	0.576581	0.608436	0.693190	0.622899	0.677538	0.537957	0.670049	0.957175	0.233406	0.751786	0.496026	0.495317	0.564879	0.450211
5516	0.640841	0.659344	0.709529	0.650967	0.653690	0.554875	0.761744	0.953898	0.070610	0.685315	0.242235	0.467265	0.275733	0.522768
5517	0.662790	0.623364	0.795127	0.712716	0.648206	0.599529	0.824451	0.970116	0.078478	0.774741	0.268907	0.524757	0.296266	0.412077

Figure 9: Data set features

4.8.2 Tools used to study the data and build the model

Jupyter Notebook is an open-source web-based interactive computing platform that is widely used in the field of machine learning and data science. It allows users to create and share documents that contain live code, equations, visualizations, and narrative text, making it an ideal tool for exploring and documenting data-driven projects.

In the context of machine learning and data science, Jupyter Notebook is often used for tasks such as data exploration and visualization, prototyping and testing machine learning models, and creating reports and presentations. It is particularly useful for working with large datasets and complex models, as it allows users to easily organize and document their work.

One of the key features of Jupyter Notebook is its ability to mix code, visualizations, and text in a single document. This makes it easy to explain and document the steps involved in a machine learning or data science project, and allows others to easily reproduce and build upon the work. Jupyter Notebook is also highly extensible, with a range of third-party extensions and libraries that can be easily installed to add additional functionality.

Overall, Jupyter Notebook is an essential tool for anyone working with machine learning and data science. It is widely used in research and education, and is a key component of many popular data science platforms and environments. It is also used in industry, where it is often used to prototype and deploy machine learning models in production [13].

Librosa is a python library that is utilized extensively for the processing and analysis of audio signals. It is particularly beneficial in constructing recommendation systems as it has the ability to extract unique features from audio signals, thus simplifying the determination of similarities between audio signals. The mel-spectrogram, for instance, is a feature that can present a visual representation of the audio signal, encapsulating both its harmonic and percussive aspects. This information can be utilized to recommend audio signals that are similar to a user's listening history.

Additionally, Librosa offers functions for examining the extracted features and determining their relationship with the audio signal. This information can be utilized to optimize the feature extraction process and determine which features are crucial for the recommendation system. The extracted features can also be utilized to train machine learning models, such as a neural network, to suggest audio signals based on a user's listening history[14].

4.8.3 build the model

Once the features were extracted, the next step was to perform data analysis to choose the best algorithm for building the recommendation system. The data was first visualized using scatter plots and histograms to gain an understanding of the distribution of the features. Based on the distribution of the features, it was decided to use the k-means algorithm for building the recommendation system. K-means is a clustering algorithm that groups similar data points together into clusters. In this case, the songs are grouped based on their similarity in terms of the extracted features.

Elbow Points Analysis:
The k-means algorithm works by first randomly selecting k initial centroids, where k is the number of clusters desired. The algorithm then calculates the distance between each data point and each centroid, and assigns each data point to the closest centroid. The centroids are then re-calculated based on the mean of the data points in each cluster, and the process is repeated until the centroids no longer change [15].

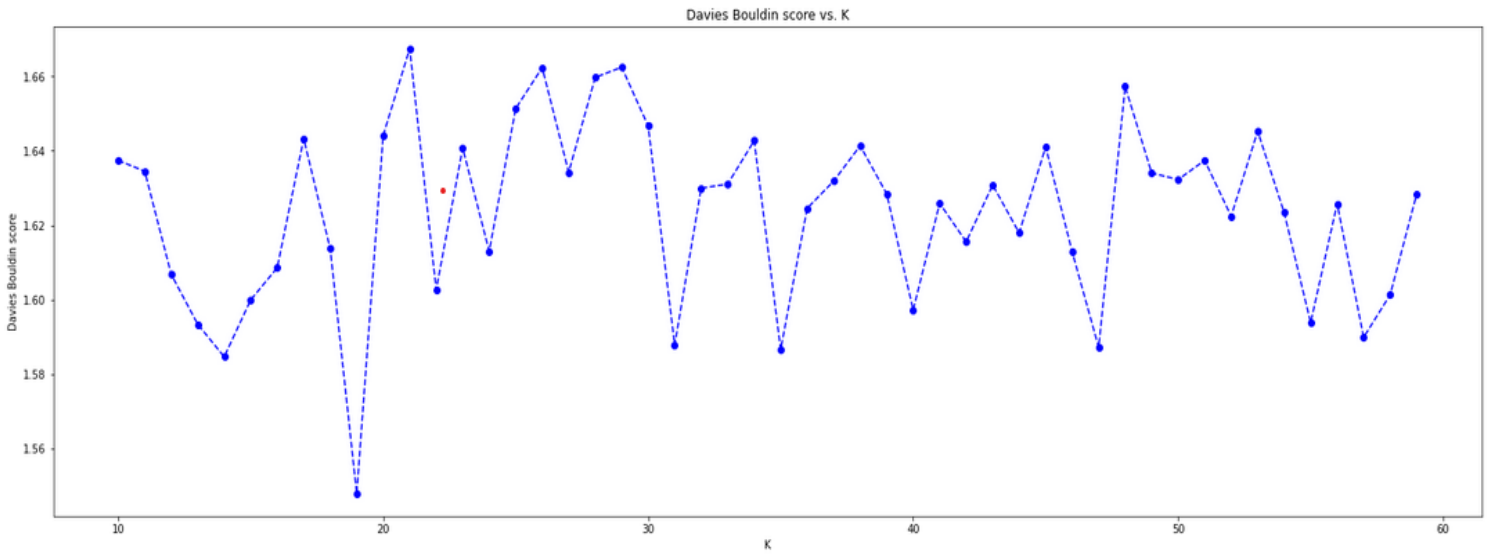


Figure 10: Davies Bouldin score vs k plot

The Davies-Bouldin Score (DBS) is a clustering validation metric that measures the quality of clustering. The DBS score is calculated by taking the average similarity between each cluster and its most similar cluster, and then dividing it by the distance between the centers of the two clusters. It is used to find the optimal number of clusters in a dataset. The K in "Davies Bouldin score vs. K" refers to the number of clusters in the dataset.

The optimal value of K is the one where the DBS score is the lowest, which indicates that the clusters are well separated and compact. Which is in our dataset is in range of 20 clusters

So we will use shows the relationship between the number of clusters and the inertia (within-cluster sum of squares) of a clustering algorithm. To get elbow point

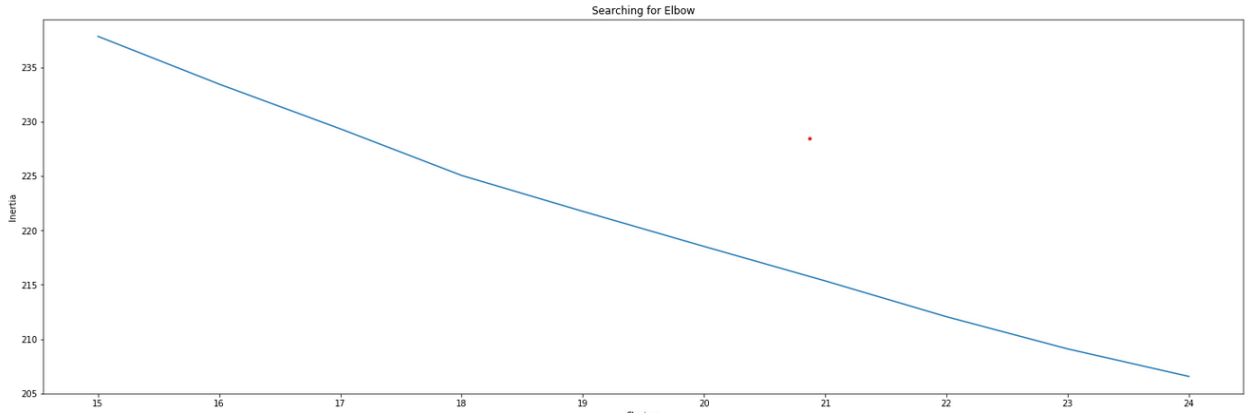


Figure 11: Elbow plot

In this specific case, the x-axis of the plot ranges from 15 to 25, and the y-values are taken from the "clusters" variable. Therefore, the plot will show how the inertia of the clustering algorithm changes as the number of clusters increases from 15 to 25.

By analyzing the plot, we can determine the optimal number of clusters for your dataset based on the position of the elbow point. This information can be used to improve the quality of your clustering results.

So, by grouping two chart we can specifically know the number of K

The elbow plot will show a curve of the inertia score for each number of clusters ranging from 10 to 45. The optimal number of clusters can be estimated from the elbow point where the curve starts to flatten out. In addition to the elbow point, the plot will also show the time taken to fit the model for each K value. This information used to evaluate the trade-off between the quality of the clustering and the computational cost of fitting the model.

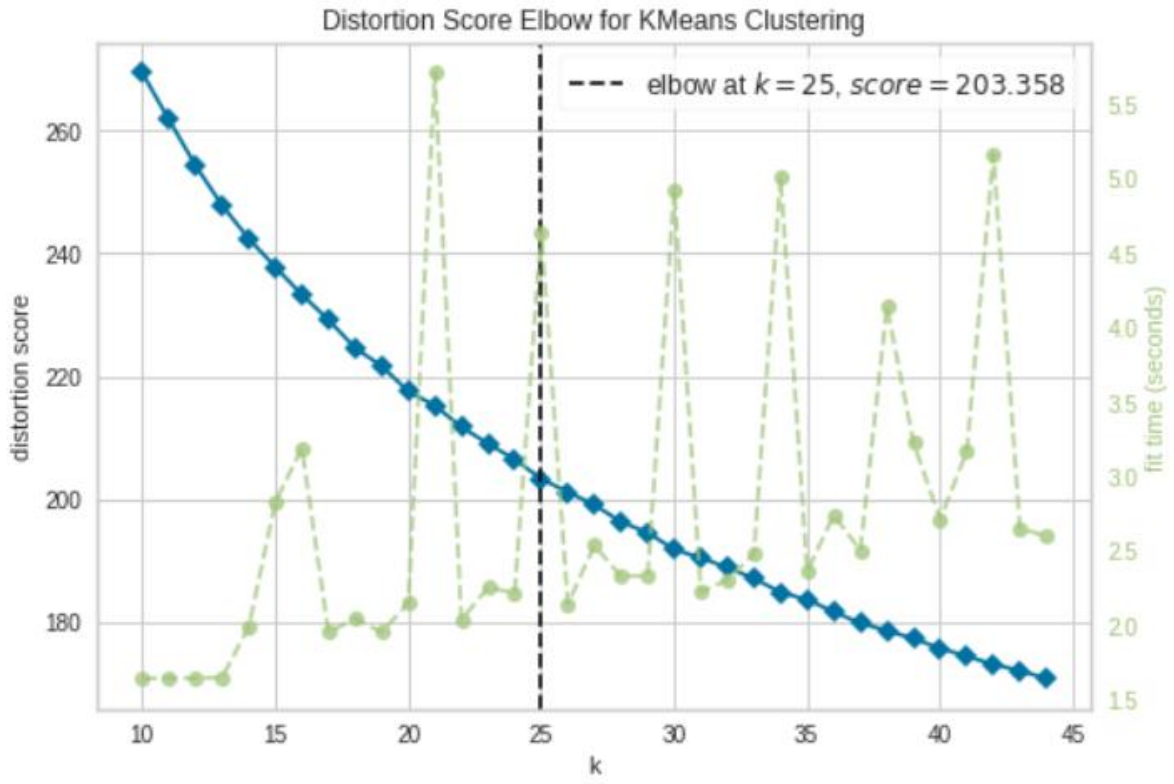


Figure 12: Distortion score elbow for k-means clustering plot

So, the best value of k is 25 cluster

By grouping similar songs together, the k-means algorithm enables us to recommend songs that are similar to a given song. For example, if a user is listening to a song with a certain set of features, the algorithm can recommend other songs with similar features.

In additional we study another algorithm such as

- Hierarchical Clustering
- hybrid approaches

By testing these three algorithms then we decided to build the model using K-means

```
In [ ]: import scipy.cluster.hierarchy as shc
        from matplotlib import pyplot
        pyplot.figure(figsize=(10, 7))
        pyplot.title("Dendrograms")
        dend = shc.dendrogram(shc.linkage(trainingDfScaled, method='ward'))
```

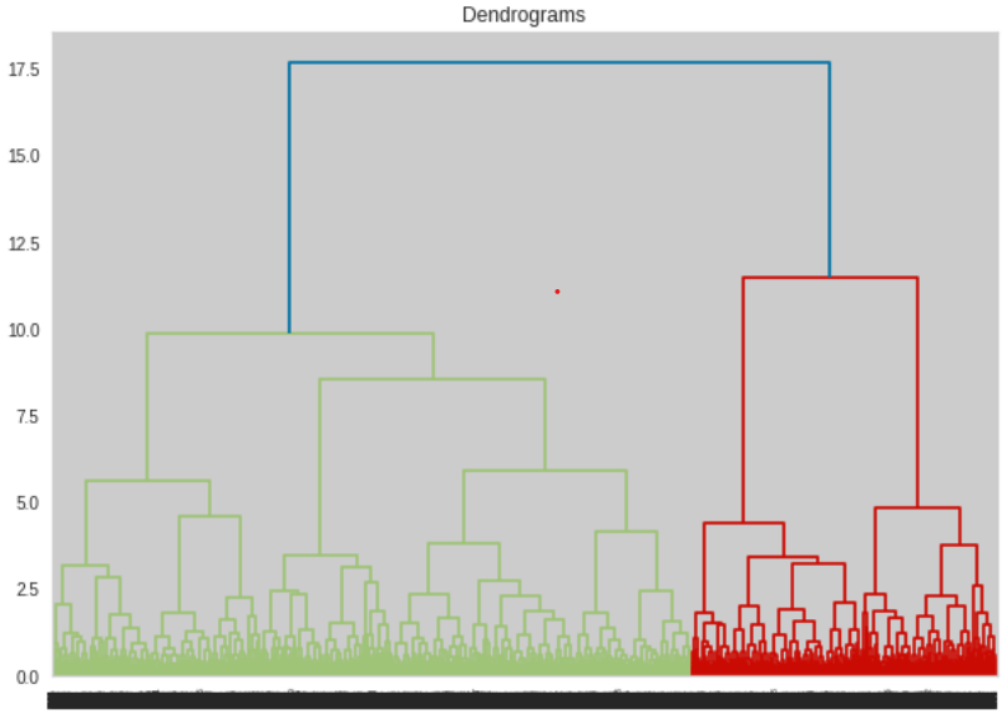


Figure 13: Hierarchical Clustering plot

4.8.4 Ranking - cosine Similarity

In machine learning, ranking using cosine similarity is a technique used to measure the similarity between two vectors. In the context of recommendation systems, it can be used to identify similar items to recommend to a user based on their past behavior and preferences.

When a new record is inserted into a recommendation system, the system can use k-means clustering to group the new record with other similar records based on their feature vectors. To do this, the system first creates a feature vector for the new record and compares it to the existing records in the system using cosine similarity. The records that are most similar to the new record are then grouped together using k-means clustering.

Based on the similarity scores, the system can then recommend items that are similar to the new record. For example, if a user has just listened to a new song in a music streaming service, the system can use cosine similarity to recommend other songs that are similar in genre, tempo, or mood.

One advantage of using K-means with cosine similarity in recommendation systems is that it is computationally efficient and can handle high-dimensional feature vectors, making it suitable for large-scale systems with many items and users. Additionally, cosine similarity is a popular method in recommendation systems because it is relatively simple to understand and implement.

It’s else use to get number of songs in the same cluster bead on similarity

By using this technique, the system handles every song uploaded by artist and put it in right cluster as following:

```
In [ ]: rankingDf
        rankingDf.to_csv("clusteredData.csv", index=False)
```

Out[83]:

roma_cqt	chroma_cens	melspectrogram	mfcc	poly_features	tonnetz	rms	zero_crossing_rate	path	artist_name	kmeansClusters25	so
0.727643	0.935546	0.161840	0.768913	0.351149	0.309264	0.411359	0.397000	folder5/Khaoula hsein - Hob Kbir (EXCLUSIVE L...	Khaoula hsein	18	hs (E)
0.724205	0.958179	0.185064	0.727174	0.423799	0.456993	0.487602	0.451821	folder5/Ibrahem Al Amer – YI3n Al Heniya (Excl...	Ibrahem Al Amer	18	Il Ar (
0.670970	0.904126	0.030395	0.635866	0.137519	0.323837	0.169175	0.370741	folder5/Inkonnu - Inko (Officiel Music Video ...	Inkonnu	14	Ink Mu
0.763156	0.943765	0.048290	0.740973	0.168955	0.424954	0.202896	0.435395	folder5/Klay ft. Sanfara - E7tilel | إحتلل [p...	Klay ft	13	[p6:

Figure 14: new songs feature uploaded to dataset and chose k

4.9 Application Interfaces

This part section the appearance of the web site Interfaces and some of the features that are available in the system.

4.9.1 Signing Up and Logging In

This section pertains to the visual appearance of the application's pages, including the login and registration screens. In particular, the sign-in page, as depicted in Figure below, allows users to enter their email and password. Upon successful authentication, the user gains access to the application.

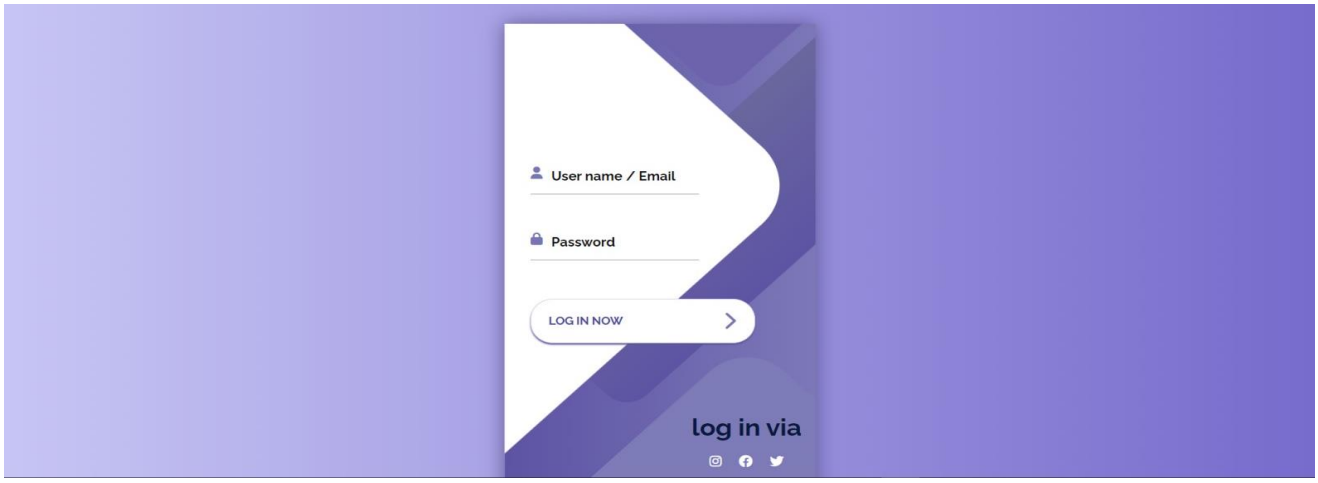


Figure 15: Login page

In the event that the user does not already have an account and wishes to create a new one, they can do so by clicking the "Register" button. This will direct them to the Sign-Up Page (also known as the "Register Page"), as depicted in Figure below. On this page, the user is required to enter accurate personal details such as their full name, email address, date of birth, and password

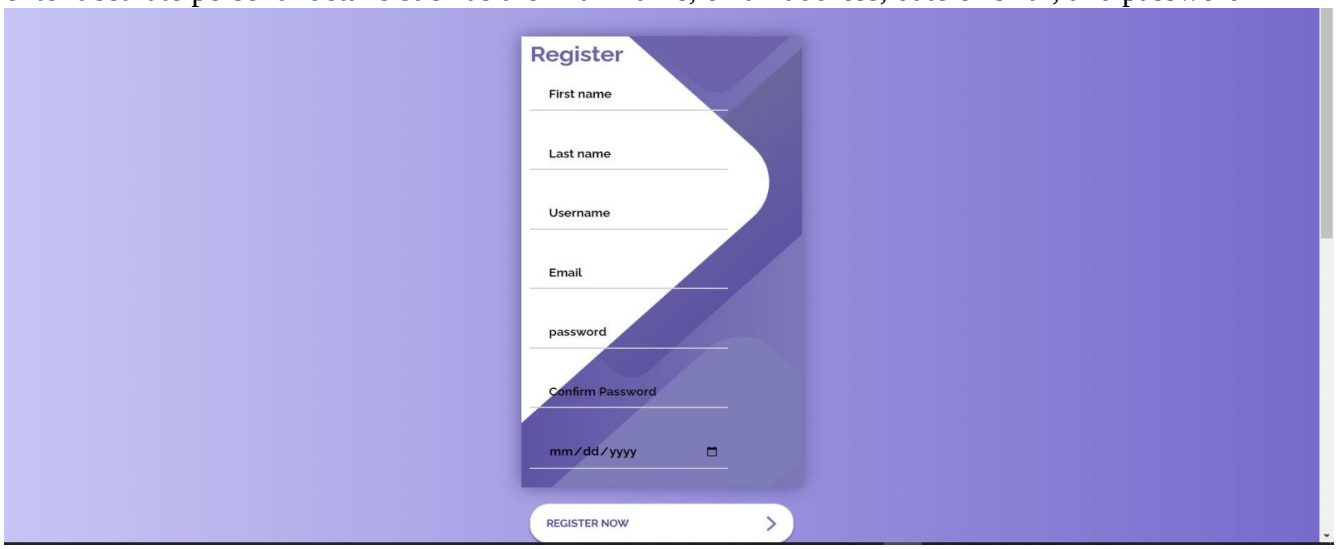


Figure 16: Register page

4.9.2 Home Page

In this part of site, any visitor can see our lists, artists and playlists if the user not login the system don’t play songs and Forword to login page

For system users

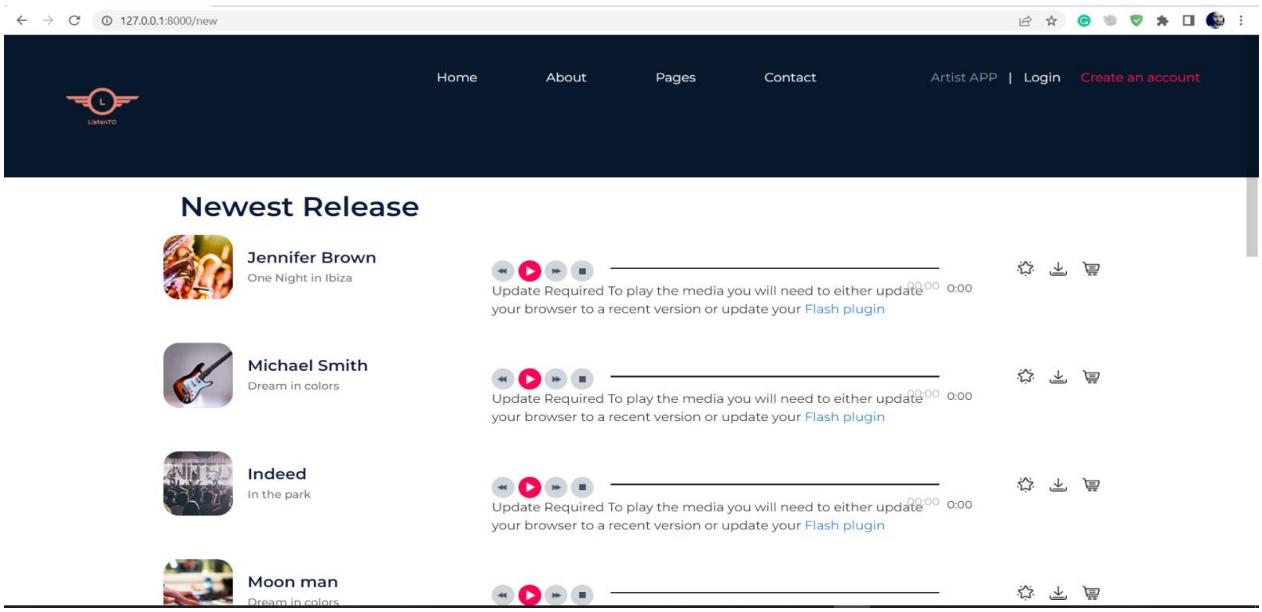


Figure 17: normal system user home page

For user how don’t login

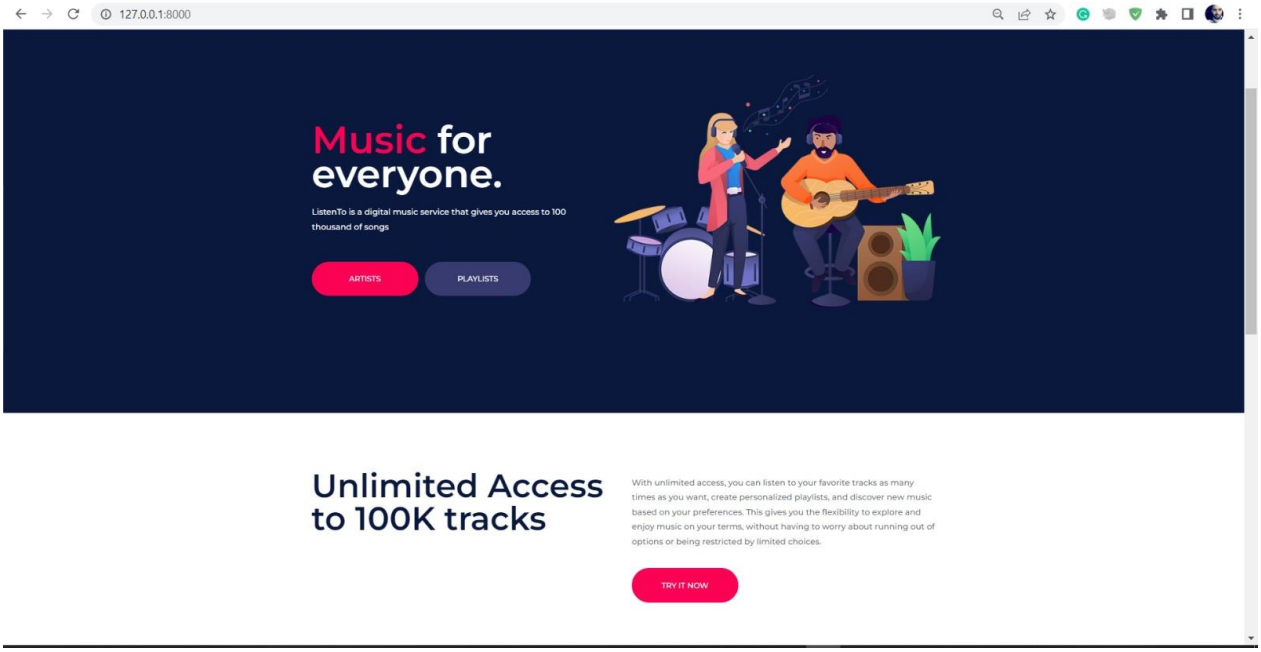


Figure 18: site home page

Any user can send to the admins of the site a massage, complaint or suggestions to take care by admins

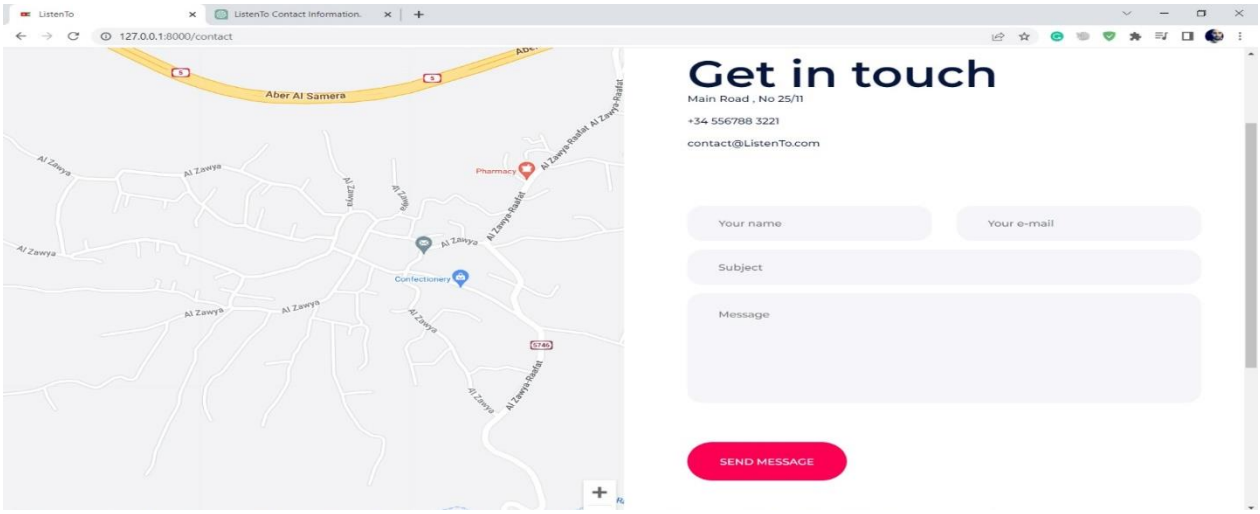


Figure 19: connect us page

4.9.3 Admin Page

The admin control panel page that can see and edit the users , accept artists request And see all events in the system and edit it and see Financial record

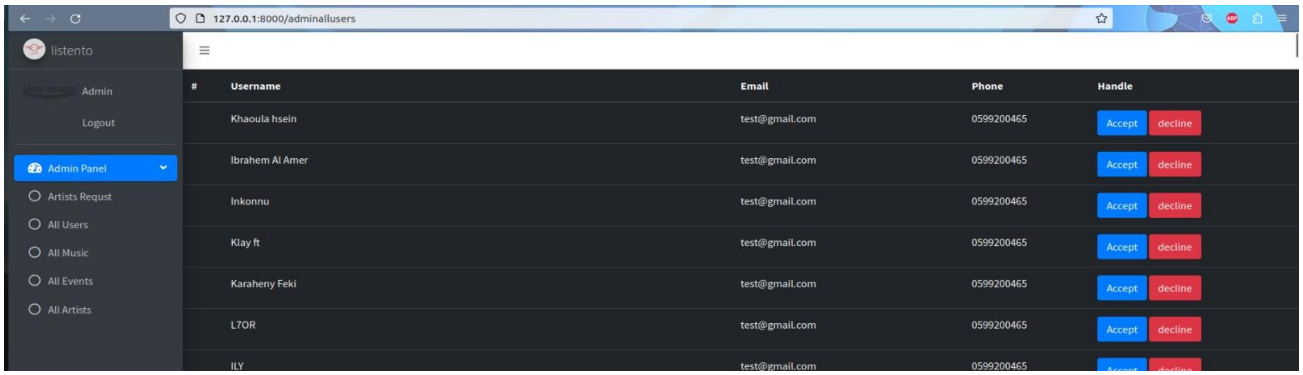


Figure 20: Admin control panel 1

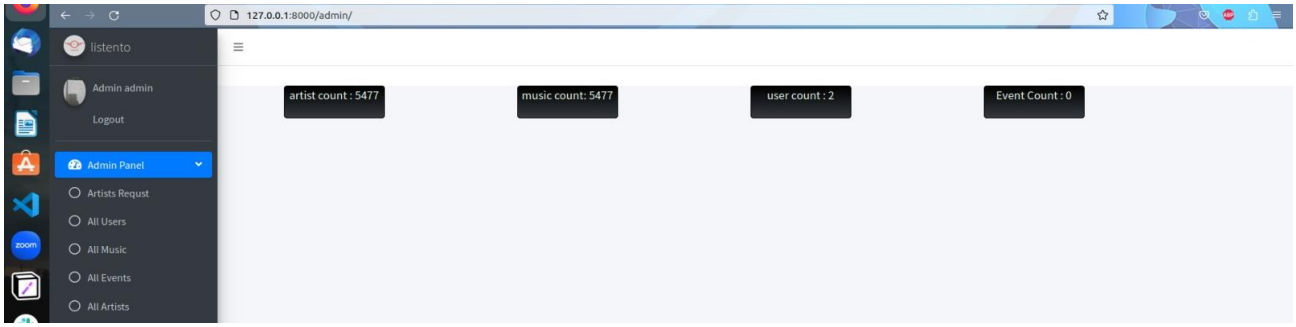


Figure 21: Admin control panel 2

4.9.4 Artist Page

The artist first register to the system as normal user then apply a form request to the admin to be an artist in the system

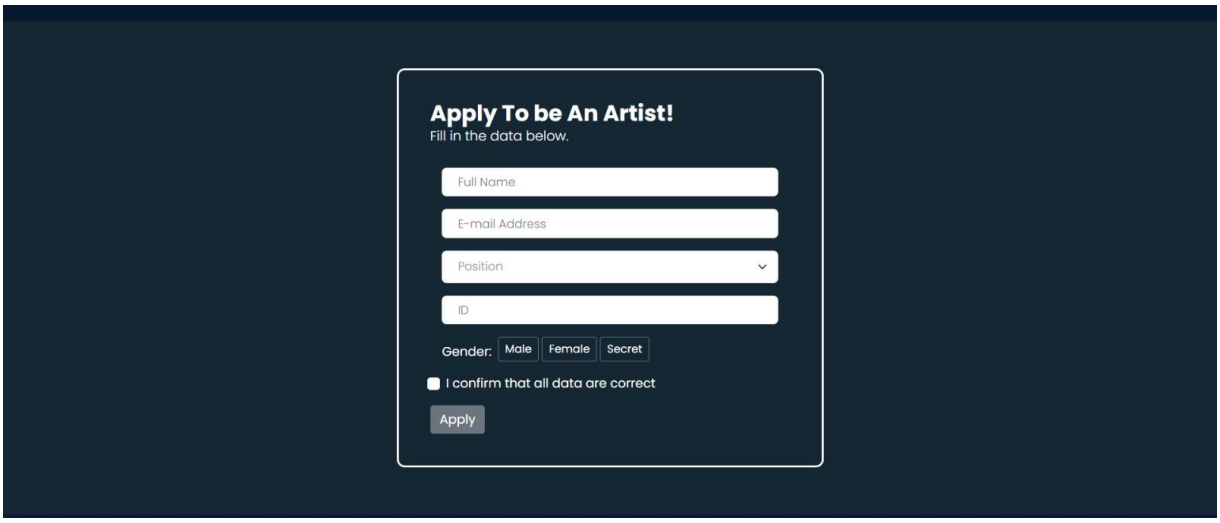


Figure 22: Apply to get artist account

The artist create the events from which contain data and ticket price and number of tickets depends on the stage name

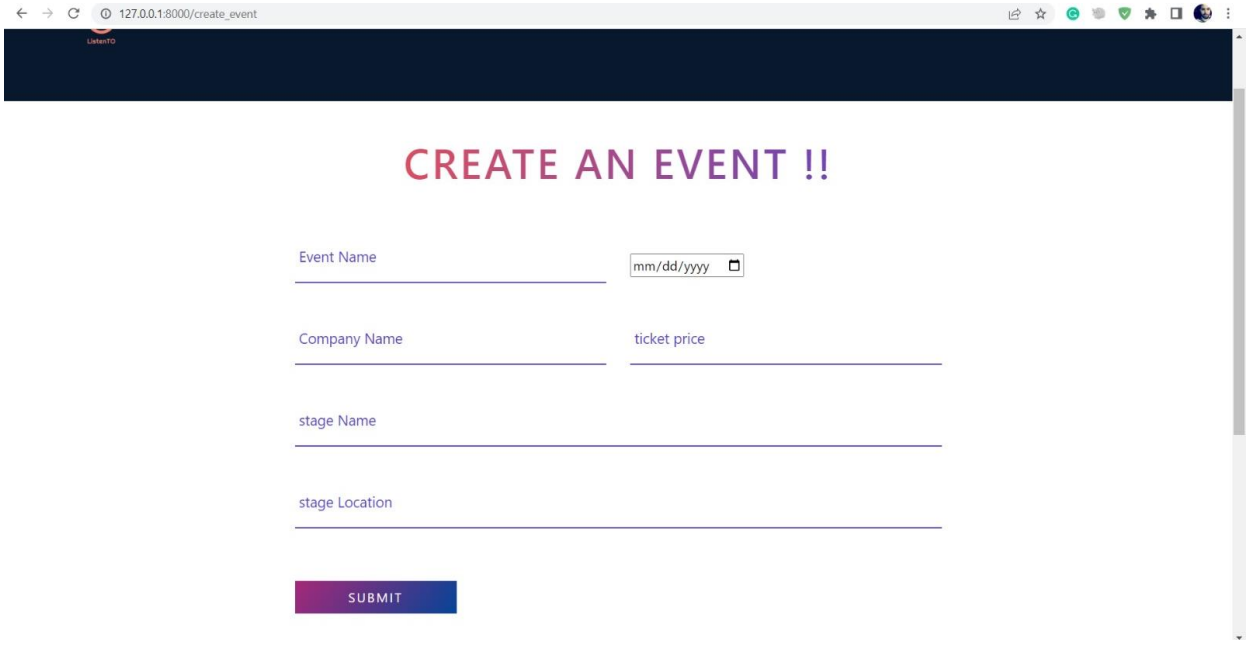


Figure 23: Artist applies to an event

Profile page for the artist which contain the information of the user , create event button and all songs with their rating

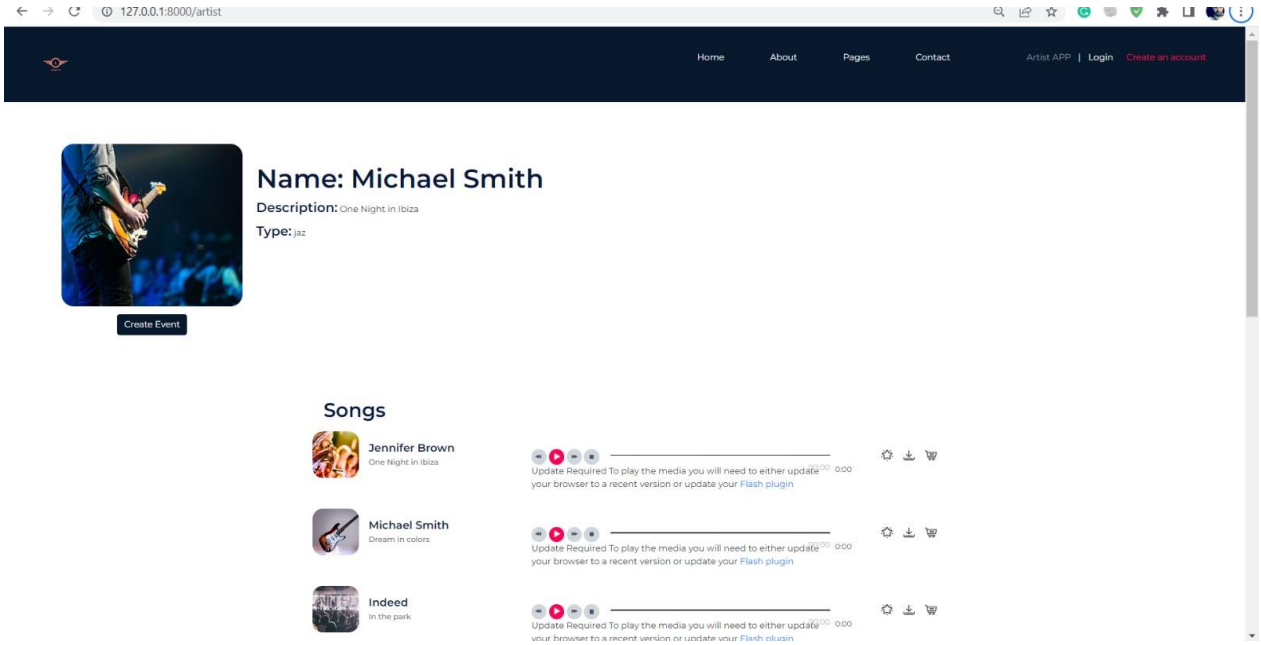


Figure 24: Artist profile page

4.9.5 User playlist

The System users create their one's playlist and add any songs to it by its name

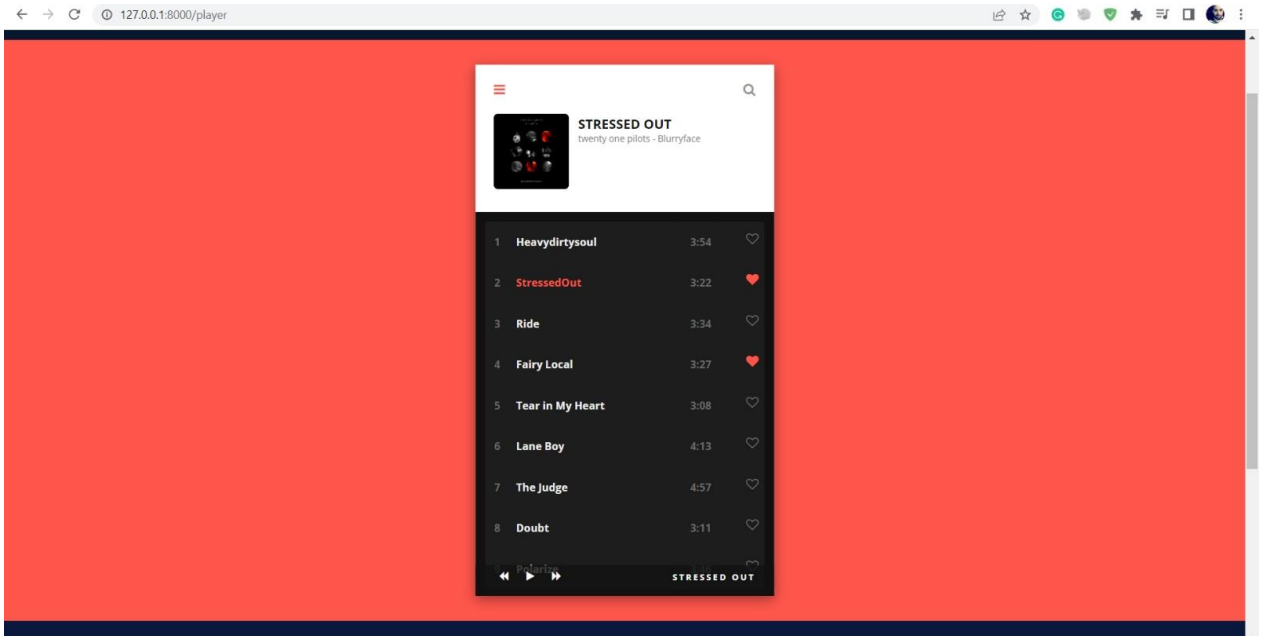


Figure 25: user's playlist

How song play in user playlist

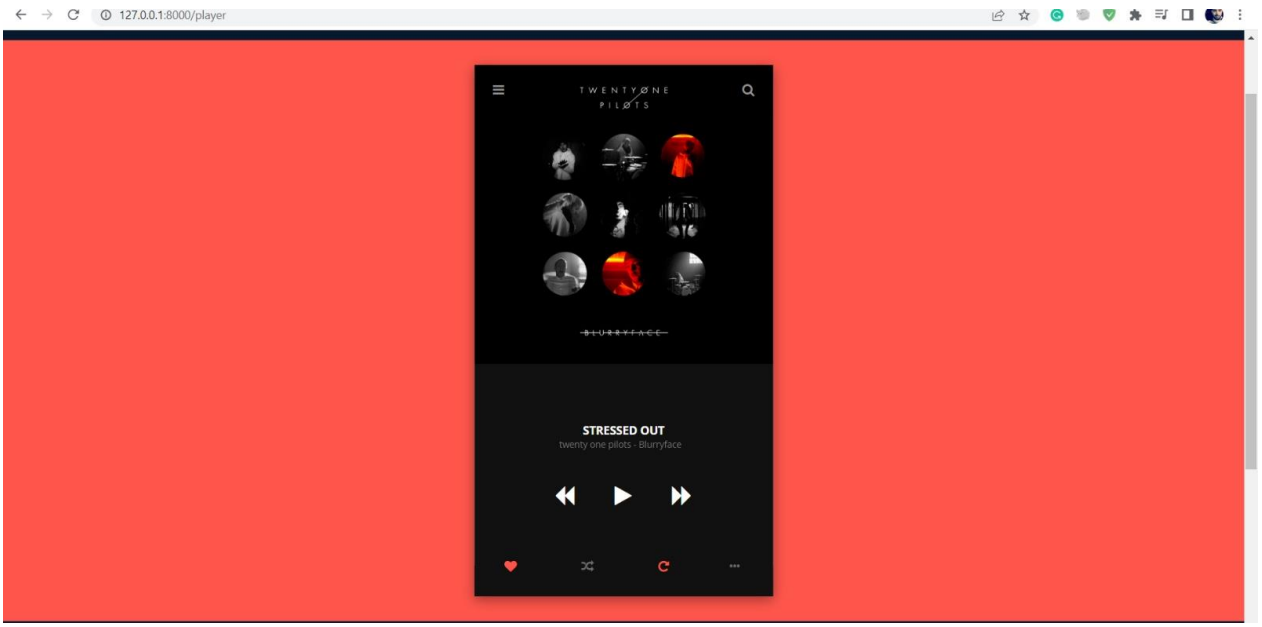


Figure 26: Songs player

Chapter 5: Conclusion and Future Work

Ultimately, we will explicate the conclusion derived from the thorough investigation and diligent execution of our study and project. Subsequently, we will provide a succinct overview of our aspirations for augmenting this project in the future.

5.1 Conclusion

In conclusion, LitenTo is an interactive music platform that utilizes AI technology to cater to each user's individual needs for free and without collection user data. With a vast collection of tens of thousands of songs, the platform also offers an event application for selling tickets

5.2 Future work:

In the near future, given enough time, there are many changes might be made to make the system more practical in real-life circumstances for the users.

- To enhance the quality of recommendations made by a model, it is often beneficial to incorporate user listening history data. This can be achieved through the utilization of machine learning techniques such as Recurrent Neural Networks (RNNs), which are capable of processing sequential data such as a user's history of listened music. However, to optimize the performance of such models, it is essential to carefully tune the hyperparameters of the model. This involves a thorough analysis of a significant amount of data to determine the optimal values of the hyperparameters, leading to a model with improved accuracy in making recommendations.
- Artist can create a live event in the site
- When the artist create an event, it's can be interactive platform, the user can like and rate or comment on a song based on the event
- Collection more .mp3 songs to create a big-data set

References

[1] https://tinuiti.com/blog/performance-display/spotify-algorithm/#:~:text=Powered%20by%20AI%2C%20Spotify's%20algorithm,features%2C%20and%20past%20listening%20habits

[2] https://talks.anghami.com/using-word2vec-for-music-recommendations/

[3] https://routenote.com/blog/is-your-deezer-flow-repetitive-fix/#:~:text=How%20does%20Deezer%20Flow%20work,get%20better%20as%20a%20result

[4] https://azure.microsoft.com/en-us/resources/cloud-computing-dictionary/what-is-java-spring-boot/

[5] https://dotnet.microsoft.com/en-us/learn/aspnet/what-is-aspnet

[6] https://mdevelopers.com/blog/a-brief-guide-through-laravel

[7] https://developer.mozilla.org/en-US/docs/Learn/Server-side/Django/Introduction

[8] https://www.talend.com/resources/what-is-mysql/

[9] "Activity," [Online]. Available: https://www.geeksforgeeks.org/unified-modeling-language-uml-activity-diagrams/#:~:text=An%20activity%20diagram%20is%20a,the%20activity%20is%20being%20executed

[10]https://www.sciencedirect.com/science/article/pii/S1875389212006220

[11]https://www.sciencedirect.com/science/article/abs/pii/S0957417406004076

[12]https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=35a87b51f7441a87adee91e12eb4d22cd2565556

[13]https://ieeexplore.ieee.org/abstract/document/7991618

[14]https://librosa.org/doc/latest/index.html

[15] https://jwcn-eurasipjournals.springeropen.com/articles/10.1186/s13638-021-01910-w

[16]https://ieeexplore.ieee.org/abstract/document/6481039

Appendices

- **website source code repository**
  https://github.com/top1million/GradProject

- **Data analysis and ML code**
  https://github.com/QossayZeineddin/Data-analysis-and-ML-for-graduate-project.git