Predictive Analytics Lab Project

School of Computer Science
University of Petroleum and Energy Studies, 248007
2024

Synopsis on
Predictive Analysis on the Influence of Music on Mental Health

Submitted by:

Name	SAP ID	Roll No.	Branch
Pranit Abraham Thomas	500100820	R2142220130	AI/ML (Hons.)
Pushp Prakhar Bhardwaj	500105429	R2142220136	AI/ML (Hons.)

Under the guidance of
Ms. Achala Shakya

Table of Contents

S. NO	Content	Page No.
1.	Abstract	2
2.	Introduction	3-4
3.	Literature Review	4-5
4.	Problem Statement	5-6
5.	Objectives	6
6.	Methodology	6-8
8.	System Requirements	8-9
9.	References	10

Abstract

Mental health and well-being are deeply intertwined with various factors, including personal experiences and external influences such as music. However, the relationship between music preferences and mental health indicators is often overlooked due to the lack of comprehensive data analysis in this area. Traditional methods of assessing mental health typically do not consider the nuances of individual experiences related to music, leading to a gap in understanding how these factors affect well-being.

This project proposes the development of a predictive model that analyzes the relationship between music preferences and mental health outcomes. By utilizing a diverse dataset, this model aims to identify patterns and correlations between individuals' music choices and psychological parameters like 'Conscientiousness', 'Neuroticism', Extraversion, K-10 scores and other personality traits. Users will be able to input their personal data regarding music preferences, lifestyle, and mental health status, and receive tailored predictions related to their mental health indicators.

Through machine learning techniques, the model not only aims to provide insights into how music can influence mental health but also to empower individuals to make informed decisions about their music consumption. By highlighting the role of music in mental well-being, this project seeks to contribute to a more holistic understanding of mental health and promote the potential benefits of music as a therapeutic tool.

Keywords: Mental Health, Music Preferences, Predictive Modeling, Machine Learning, Well-being

1. Introduction

In today's fast-paced world, mental health concerns are on the rise, with an estimated 1 in 8 people worldwide living with a mental disorder. According to the World Health Organization (WHO), mental health conditions are projected to account for 20% of the global disease burden by 2030, with depression and anxiety among the leading causes of disability. Music, as a universal language, has long been associated with emotional and psychological regulation, offering both therapeutic and expressive benefits. Research has shown that music can influence mood, reduce stress, and provide comfort during challenging times. With this understanding, our project seeks to develop a platform that assesses the relationship between music preferences and mental health indicators. By creating a model that predicts key psychological parameters—'Unhealthy', 'Healthy', 'K-10', 'Openness', 'Conscientiousness', 'Extraversion', 'Agreeableness', and 'Neuroticism'—we aim to provide valuable insights into how music affects well-being and personality traits.

Our project is built on an extensive dataset that captures a wide range of variables related to demographics, music habits, and mental health. The dataset consists of over 60 columns, including inputs such as 'Gender', 'Age', 'Nationality', 'Your interest in music?', and 'What kind of music do you like to listen to when you feel low?'. These inputs are essential in understanding the user’s interaction with music and its potential impact on their mental state. For instance, questions related to emotional responses to music—such as 'I hide in my music because nobody understands me, and it blocks people out' and 'Music helps me relax'—offer insight into how music is used as a coping mechanism. Additionally, the dataset includes mental health-related questions based on the Kessler Psychological Distress Scale (K-10), a well-established tool for measuring anxiety and depression symptoms, which forms the foundation for predicting the 'K-10' score.

The model we are developing leverages machine learning techniques to analyze user inputs and predict outcomes for the aforementioned psychological parameters. 'Openness', 'Conscientiousness', 'Extraversion', 'Agreeableness', and 'Neuroticism' are key personality traits that are critical in understanding an individual's behavior, relationships, and emotional resilience. For instance, high levels of 'Extraversion' might indicate a tendency to engage in social activities, while 'Neuroticism' often correlates with emotional instability and higher stress levels. By offering personalized predictions based on these traits, the system provides users with a deeper understanding of their psychological landscape. It not only shows how their music habits relate to mental health but also provides actionable insights into areas that may require attention. For example, users who score higher on 'Neuroticism' and 'Unhealthy' might be encouraged to adopt healthier music practices to improve their mental well-being. Similarly, individuals with lower 'Conscientiousness' scores may discover how music influences their ability to focus and organize.

The potential applications of this project extend beyond individual users to mental health professionals and researchers. As mental health awareness continues to grow, tools like ours could offer a novel approach to understanding emotional well-being through the lens of music therapy. Studies have already shown that music therapy can be an effective treatment for anxiety, depression, and other mental health conditions. By combining this with a predictive model that analyzes a user's personality and mental health status, we can provide personalized music-based interventions that enhance mental well-being. For example, an individual experiencing high levels of stress or anxiety could receive recommendations on specific types of music that may help alleviate their symptoms. Furthermore, the platform could be integrated into wellness apps, allowing users to track their mental health progress over time and adjust their music habits accordingly.

In conclusion, this project represents a significant step towards understanding the intricate connection between music and mental health. By predicting parameters such as 'Unhealthy', 'Healthy', 'K-10', 'Openness', 'Conscientiousness', 'Extraversion', 'Agreeableness', and 'Neuroticism', we aim to offer users valuable insights into how their music choices shape their emotional and psychological states. As mental health continues to be a critical issue globally, our platform has the potential to become a valuable tool for both individuals and professionals looking for personalized ways to improve well-being. Through the power of machine learning and the universal influence of music, this project aspires to contribute to a deeper understanding of mental health, providing a scalable and innovative solution for those seeking to enhance their mental and emotional well-being.

2. Literature Review

The relationship between music genres and mental health has garnered increasing attention in recent research, particularly due to music's potential therapeutic effects. Two key studies explore this connection: Rahman et al.'s study on using music for mental health therapy through machine learning, and McFerran et al.'s critical synthesis on music's impact on adolescent mental health.

Music Therapy and Mental Health Care

Rahman et al. (2021) delve into the application of music therapy for mental health care by analyzing physiological signals from participants listening to various music genres. They highlight that music evokes diverse emotions, reflected in physiological signals such as Electrodermal Activity (EDA) and Blood Volume Pulse (BVP). The study reveals that neural network models achieved a classification accuracy of 99.2% in distinguishing between genres based on these signals, further showing that classical and instrumental music, often used in therapy, can help reduce stress, anxiety, and even improve sleep quality [1]. The paper underscores that different genres can influence emotional states, with a long-term goal of reducing musicogenic epilepsy. This study reinforces the idea that music genres hold distinct emotional power, a premise crucial for understanding music's therapeutic potential.

Music and Adolescent Mental Health

McFerran et al. (2013) take a different approach by focusing on how music engagement affects adolescent mental health, specifically concerning depression. They performed a critical synthesis of 33 studies, analyzing the relationship between musical preferences and mental health outcomes. A notable finding is the correlation between preferences for certain genres, such as heavy metal and rap, and negative health outcomes like depression and social alienation. On the other hand, engaging with music through creation, such as playing instruments or participating in group music-making, was associated with enhanced social skills, emotional expression, and reduced anxiety [2]. The study emphasizes that adolescents' musical behaviors, not just their preferences, play a crucial role in influencing mental health outcomes.

Comparison and Synthesis

Both studies highlight that the effects of music on mental health are genre-dependent. Rahman et al. focus more on the physiological response to different genres, suggesting that classical and instrumental music can be powerful tools for therapy. Meanwhile, McFerran et al. concentrate on the behavioral aspect of music engagement among adolescents, finding that music-making is generally beneficial, while certain genres of music listening can be associated with negative mental health outcomes like depression. Together, these studies suggest a multi-faceted approach to understanding the connection between music and mental health: it is not only the genre but also the mode of engagement that plays a significant role.

In conclusion, these studies provide a robust foundation for exploring how music genres and engagement modes can predict mental health outcomes, and both support the idea that music can be a powerful tool in managing emotional well-being.

3. Problem Statement

Mental health issues are increasingly prevalent, yet there is a lack of effective tools to predict mental well-being based on personal habits and preferences, especially related to music. Music often serves as a coping mechanism, but the relationship between music preferences and mental health indicators—such as 'Unhealthy', 'Healthy', 'K-10', 'Openness', 'Conscientiousness', 'Extraversion', 'Agreeableness', and 'Neuroticism'—remains underexplored. This lack of clarity leaves individuals and professionals without actionable insights into how music impacts mental health.

To bridge this gap, a system that collects user data on music habits and predicts mental health outcomes using machine learning is needed. Such a platform would offer users personalized insights into their mental well-being and provide professionals with a valuable tool to tailor treatments based on music preferences. Developing this solution would significantly improve the understanding of music's role in mental health, offering a structured, data-driven approach to better manage emotional and psychological well-being.

4. Objectives

- To clean and preprocess the dataset, including encoding categorical columns and handling missing data, ensuring it is ready for model training.
- To split the dataset into training, testing, and validation sets, and train a machine learning model to predict mental health and personality outcomes.
- To enable users to input their own responses and predict possible scores for the target columns: 'Unhealthy', 'Healthy', 'K-10', 'Openness', 'Conscientiousness', 'Extraversion', 'Agreeableness', and 'Neuroticism'.

5. Methodology

1) Data Preparation and Preprocessing

- Data Cleaning: Handle missing data, remove duplicates, and perform necessary data transformations.
- Encoding Categorical Variables: Apply appropriate encoding techniques to convert categorical columns into numerical representations.
- Feature Scaling: Standardize or normalize features as needed for optimal model performance.

2) Model Development

- Define Objectives: Set clear goals for predicting target variables related to mental health and personality traits, including 'Unhealthy', 'Healthy', 'K-10', 'Openness', 'Conscientiousness', 'Extraversion', 'Agreeableness', and 'Neuroticism'.
- Model Selection: Choose machine learning models suitable for the task, such as regression or classification models.
- Data Splitting: Divide the dataset into training, testing, and validation sets for model training and evaluation.

3) Model Training and Testing
- Training: Train the selected machine learning model using the training data and optimize hyperparameters where necessary.
- Testing and Validation: Evaluate model performance on the test and validation sets, fine-tune for accuracy, and ensure generalization of the model.

4) User Interaction and Prediction Interface
- User Input: Allow users to input their own responses for each column, such as demographic data, music preferences, and emotional indicators.
- Prediction Output: Enable real-time predictions for the target columns based on user inputs, providing them with scores for mental health and personality traits.

5) Integration and Testing
- System Integration: Combine model predictions with the user interface, ensuring smooth interaction and real-time responses.
- Unit Testing: Test individual components, including input validation, model predictions, and data handling.
- System Testing: Conduct end-to-end testing to verify that the model correctly predicts and outputs scores based on user inputs.

6) Documentation
- Technical Documentation: Record the technical details of the model, data preprocessing methods, and system architecture.
- User Documentation: Create a user guide to help users understand how to input data and interpret their results.

7. System Requirements

In developing the model to predict the ‘Influence of Music on Mental Health’, a well-rounded approach to software requirements is essential to ensure the platform's accuracy, usability, and scalability. The system must efficiently collect and process user data, apply machine learning models, and provide insightful predictions on how different music genres influence mental health outcomes. This section outlines the key functional and non-functional requirements needed to create a reliable platform that can assist in understanding the psychological impacts of music and support personalized music therapy for improving mental well-being.

Functional Requirements

These requirements define what the system must do to achieve the project goals.

- **Data Collection & Preprocessing:**
  - The system must collect user data on musical preferences (genres) and mental health indicators.
- **Music Genre Classification:**
  - The system should classify music genres based on predefined categories like classical, pop, instrumental, etc.
- **Correlation Analysis:**
  - The system should allow for predictive analysis between music genres and mental health using data such as user feedback.
- **Data Encoding and Scaling:**
  - The system shall utilize joblib to save the encoding schemes applied to each input column individually. This ensures that the same transformation can be consistently applied to new data inputs, maintaining the integrity of the data processing pipeline.
- **Standard Scaling:**
  - The system shall implement a standard scaler to normalize numerical features in the dataset. The scaler will be saved using joblib to facilitate seamless integration in future data processing and model prediction stages.
- **Model Persistence:**
  - The trained machine learning model shall be saved using joblib, allowing for efficient retrieval and deployment without the need for retraining. This enables the system to provide predictions in real-time based on the saved model.
- **XGBoost Implementation:**
  - The system shall utilize the XGBoost algorithm for predictive modeling due to its effectiveness in handling complex datasets and high dimensionality. XGBoost's gradient boosting framework will be employed to train the model, optimizing performance and accuracy in predicting key psychological parameters such as 'Unhealthy', 'Healthy', 'K-10', 'Openness', 'Conscientiousness', 'Extraversion', 'Agreeableness', and 'Neuroticism'.

Non-Functional Requirements

These define the quality attributes of the system, including performance, scalability.

- **Performance:**
    - The system should process large datasets (music data, user data) efficiently, ensuring minimal lag during analysis.
    - Machine learning models should predict outcomes with reasonable response times, typically under a few seconds.
- **Scalability:**
    - The system should be scalable to handle an increasing number of users and larger datasets.
    - It should be capable of handling the complexity of real-time physiological data collection if applied in future implementations.

8. References

1. Rahman, J. S., Gedeon, T., Caldwell, S., Jones, R., & Jin, Z. (2021). *Towards effective music therapy for mental health care using machine learning tools: Human affective reasoning and music genres*. JAISCR, 11(1), 5-20 [6].
2. McFerran, K. S., Garrido, S., & Saarikallio, S. (2013). *A critical interpretive synthesis of the literature linking music and adolescent mental health*. Youth & Society, 45(1), 1-18 [7].