Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2023 SCPE. Volume 24, Issues 4, pp. 919–939, DOI 10.12694/scpe.v24i4.2467

CONCOLLA – A SMART EMOTION-BASED MUSIC RECOMMENDATION SYSTEM FOR DRIVERS

JIGNA PATEL,* ALI ASGAR PADARIA,† ARYAN MEHTA,‡ AARYAN CHOKSHI,§ JITALI PATEL,¶ AND RUPAL KAPDI ‖

Abstract. Music recommender system is an area of information retrieval system that suggests customized music recommendations to users based on their previous preferences and experiences with music. While existing systems often overlook the emotional state of the driver, we propose a hybrid music recommendation system - ConCollA to provide a personalized experience based on user emotions. By incorporating facial expression recognition, ConCollA accurately identifies the driver's emotions using convolution neural network(CNN) model and suggests music tailored to their emotional state. ConCollA combines collaborative filtering, a novel content-based recommendation system named Mood Adjusted Average Similarity (MAAS), and apriori algorithm to generate personalized music recommendations. The performance of ConCollA is assessed using various evaluation parameters. The results show that proposed emotion-aware model outperforms a collaborative-based recommender system.

Key words: Emotion, mood, music, recommendation system, matrix factorization collaborative filtering, personalized content-based recommendation, deep learning, apriori algorithm, associative rule mining

1. Introduction. Music technology is advancing at a rapid pace, providing people with greater access to music than ever before. People primarily use music for self-awareness, fostering social connections, arousal, and mood control [26]. With laptops, tablets, smartphones, and other mobile devices, people can easily listen to music while working, exercising, travelling, or relaxing at home. Music has become an essential part of our daily lives, and it's hard to imagine a future without it. Listening to music while driving is a popular activity for many people. As automotive audio systems have become more sophisticated, drivers can now enjoy a more immersive listening experience. Modern cars often come equipped with state-of-the-art sound systems that offer features such as Bluetooth connectivity, HD radio, and even integrated streaming services. Furthermore, a lot of people stream music through their car's audio system using their smartphones or portable music players. Fig. 1.1 showcases that listening to music in the car[25] is the most popular among all the different places where people listen to music[9], with around 66% of listeners choosing this location. This trend can be attributed to the positive impact music has on our mood, relaxation, and concentration, making it a popular choice for many. With the convenience of linking smartphones to car audio systems, enjoying your favourite music while driving has become effortless. As a result, car music listening has become a common sight on the roads.

However, it is important for drivers to maintain their focus on the road and avoid letting their love for music distract them while driving. It is essential to remember that safe driving should always be the top priority. According to the World Health Organization's (WHO) World Report on Road Traffic Injury Prevention, there are frighteningly many road traffic fatalities globally, with 1.2 million reported each year [1]. Additionally, approximately 50 million people suffer from disabilities or injuries resulting from road accidents each year.

*Department of Computer Science and Engineering, Institute of Technology, Nirma University, Ahmedabad, Gujarat, India, (jignas.patel@nirmauni.ac.in)
†Department of Computer Science and Engineering, Institute of Technology, Nirma University, Ahmedabad, Gujarat, India (20bce010@nirmauni.ac.in).
‡Department of Computer Science and Engineering, Institute of Technology, Nirma University, Ahmedabad, Gujarat, India (20bce018@nirmauni.ac.in).
§Department of Computer Science and Engineering, Institute of Technology, Nirma University, Ahmedabad, Gujarat, India (20bce002@nirmauni.ac.in).
¶Department of Computer Science and Engineering, Institute of Technology, Nirma University, Ahmedabad, Gujarat, India (jitali.patel@nirmauni.ac.in), Corresponding Author
‖Department of Computer Science and Engineering, Institute of Technology, Nirma University, Ahmedabad, Gujarat, India (rupal.kapdi@nirmauni.ac.in)

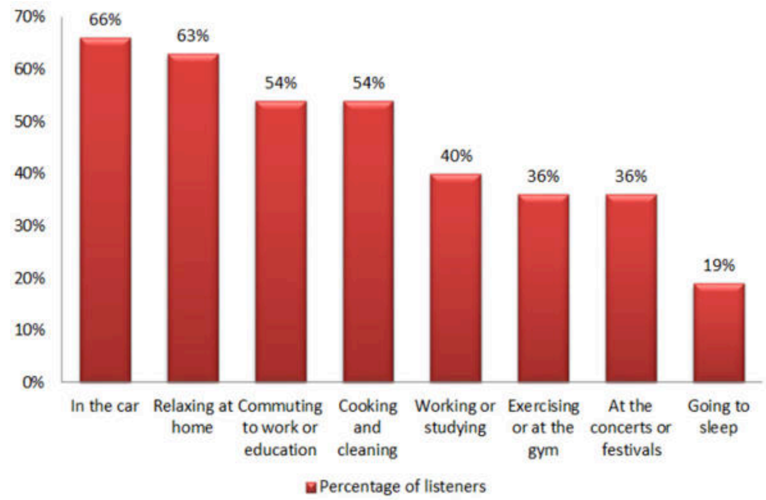

Fig. 1.1: Percentage of listeners in different environments

Analyses of European accident data have revealed that a significant proportion of road accidents, ranging from 10% to 20%, are caused by factors such as reduced driver alertness due to negative emotions or fatigue[17].

Music streaming services like Spotify have revolutionized the way we listen to music by offering vast collections of songs, personalised playlists, and intelligent recommendation systems that suggest new music based on our preferences and behaviour. However, traditional music recommendation systems often fail to capture the emotional and effective qualities of music, which are essential factors in our enjoyment and engagement with music.

In this study, we present ConCollA - an emotional-based music recommendation system for drivers. Our system successfully detects the driver's emotional state using facial expression recognition, and correspondingly recommends music. We have used a number of techniques, such as matrix factorization collaborative filtering[22], personalised content-based recommendations, and apriori algorithms, to develop these personalised suggestions[23]. This all-encompassing strategy makes sure that the music recommendations match the driver's feelings, improving their overall driving experience.

1.1. Significance of research . The conducted research in this study makes noteworthy contributions to various critical domains. Primarily, addressing the issue of road traffic accidents and their consequences assumes paramount significance due to the staggering number of deaths and injuries occurring worldwide each year. Thus, exploring innovative approaches to mitigate factors leading to accidents becomes imperative. This research endeavors to achieve precisely that by focusing on the emotional well-being of drivers through a custom music recommender system. The aim is to enhance their driving experience and potentially alleviate negative emotions or fatigue-related reduced alertness.

Moreover, ConCollA introduces a novel approach to the domain of music recommendation systems. By amalgamating content-based recommendations, association rules, and user mood/emotion information, it strives to enhance the accuracy, relevance, and personalization of music suggestions tailored for drivers. This advancement has the potential to revolutionize music recommendation and enjoyment during driving, thereby leading to more pleasurable journeys and heightened driver satisfaction.

In addition to these advancements, the incorporation of user mood and emotions into the recommendation algorithm constitutes a pivotal aspect of this research. This integration of emotional information aims to create a more comprehensive and tailored music recommendation experience. Notably, this facet of the study contributes to the growing body of research investigating the intersection of human emotions, technology, and personalized systems. By acknowledging user emotions, the recommender system can better cater to individual preferences and emotional states, thus enriching the overall driving experience.

1.2. Problem Definition. Previous studies have indicated the potential of playing relaxing music while driving to uplift the driver's mood and promote relaxation [31]. Music has proven to be a valuable tool in enhancing driving comfort and performance[7]. However, current music recommender systems for driving typically fall into one of three categories: content-based, collaborative filtering, or hybrid approaches [8]. Nevertheless, recent advancements have shown promising results in music recommenders that consider user emotions and contextual factors [4]. Researchers recognize the importance of incorporating user emotions and context to improve the personalization and effectiveness of music recommendations. By considering the driver's emotional state, driving conditions, and preferences, these innovative approaches aim to create more tailored and engaging music experiences during driving.

Despite the progress in utilizing user emotions and context in music recommenders, there is still ample room for further exploration and improvement.

In this study, we propose a novel approach for music recommender system specifically designed to cater to the unique needs of drivers. Our method synergizes the strengths of content-based and collaborative filtering techniques. Additionally, we introduce three novel enhancements to optimize the recommendation process by incorporating the driver's mood or emotions. Firstly, we present an original content-based recommender that utilizes music attributes and driver preferences to deliver personalized recommendations. Secondly, we integrate association rules into the recommendation process, enabling the identification of music tracks that are likely to resonate well with the driver's current selection or mood. Lastly, we give due consideration to user mood and emotion information as an integral factor in the recommendation algorithm, allowing us to tailor music suggestions to align with the driver's emotional state.

Through the integration of these components, ConCollA seeks to enhance the accuracy and relevance of music recommendations for drivers, ultimately elevating their driving experience and potentially mitigating negative emotions or fatigue-related reduced alertness. This research endeavors to advance the field of music recommender systems for drivers by offering an innovative method that takes into account driver mood and emotions, thus contributing to the overall goal of promoting driver well-being and road safety.

2. Related Work. Mood-based music recommendation systems have received a lot of attention in recent years because of their capacity to personalise music playlists depending on users' mood. In the context of driving, where mood and emotions can have a significant impact on the driving experience, such recommendation systems have the potential to improve driver pleasure and safety. We give a review of related work in the field of mood-based music recommendation systems in this part. A brief literature review is tabulated in Table 2.1.

The paper [5] provides an example of an approach for obtaining subjective assessments of the applicability and influence of particular contextual factors on music track ratings. The study shows that this method makes it possible to gather helpful reviews and makes it easier to create a recommender system that is aware of its environment. The predictive model, which extends Matrix Factorization (MF), is evaluated offline, and the results show significant improvements over both non-personalized and conventional personalised predictions based on MF. Additionally, a mobile application has been created to offer customers personalised and context-aware music recommendations while they are in a moving vehicle.

Context-aware music delivery systems make intelligent music choices based on contextual awareness to promote safer driving. The effectiveness and efficiency of music recommendation are two essential components of situation-aware music delivery.Real time context-based music suggestion relies on efficiency, since the music delivery system must react fast to any changes in the environment and play the appropriate music before the sensed context-data is rendered useless. In [17], the authors propose cloud and crowd-sensing based approach for music mood-mapping that can be used in context-aware music recommendation systems. This will help to speed up the music mood-mapping process and significantly increase the effectiveness of situation-aware music delivery systems.

The authors in [3] propose an emotion-aware personalized music recommendation system (EPMRS) that combines deep convolutional neural networks (DCNN) and weighted feature extraction (WFE) to establish the correlation between user data and music. The EPMRS utilizes DCNN to extract latent features from music data and WFE to generate implicit user ratings for music.The study also includes the implementation of Android and iOS apps to collect user feedback, which confirms the system's ability to reflect user preferences based on their emotions.Future work could explore the use of social media data and alternative classification techniques

Table 2.1: Literature Review

References	**Year**	**Objective**	**Methodology**	**Limitations**
Çano *et. al.* [7]	2017	Contextual mood-based music recommender system to have a positive impact on driver's mood.	A multi-point approach featuring three modules; semantic mood model to identify emotions in music, use of cardiovascular data to analyse user's emotional state and driving style recognition through data from dashboard	Few categories of moods to work on, only considers driver's heart-based data.
Deng *et. al.* [8]	2012	Building a recommender system based on the music's acoustic features and employing techniques to create an optimized recommended list of songs.	Creation of a response-arousal-valence measure to describe the emotion of a song. Recommendation done by detecting user's emotion and predicting songs likely to be heard by the user.	Model approach too complex for a low-latency application, high computational cost.
Joshi *et. al.* [16]	2021	Employing LSTM and CNN-based architectures for detection of driver's emotion.	Comparative analysis of four different model architectures, namely LSTM, CNN, LSTM-CNN and CNN-LSTM.	Does not throw light on how the detected emotions are translated into music recommendations.
Singh *et. al.* [28]	2020	Creating a hybrid approach to estimate the likelihood that a song would be heard by the user.	Using various combination of techniques such as Singular-Value Decomposition (SVD) and Factorization Machines (FM), in addition to content-based and collaborative filtering to recommend songs.	The recall shows a drop as compared to other recommendation techniques.
Lin *et. al.* [19]	2022	Personalizing music recommendations based on collaborative filtering and SVD techniques.	Combination of user-based characteristics, matching similar users and music-based filtering, matching songs based on their musical features.	Does not take into account various human-based data, eg. : mood, emotion, mental state, etc.
Iyer *et. al.* [13]	2017	Employs facical recognition techniques to recommend suitable playlists.	Using the Viola-Jones algorithm to detect critical features in the image by creating an integral image, calculating Haar-like features and using Adaboost to train the classifier	Playlist recommendation only limited to facial emotions. Does not take into account various collaborative filtering methods.
Henry *et. al.* [12]	2022	Implementation of Apriori algorithm to extract patterns in music playlists that can aid in better music recommendation.	Data preprocessing to reduce 20,000 data rows into 532, used Apriori algorithm to find associative rules between songs.	Lack of quantitative results which proscribe the effectiveness of the proposed method.
Proposed approach	2023	Recommending music to drivers based on a novel technique that combines content-based filtering in conjunction with facial emotion recognition, collaborative filtering and associative rule mining to suggest songs based on the emotion of the driver.	Implemented a fusion of CNN-based emotion recognition model and content-based filtering model that suggests music based on emotional features. The addition of a collaborative filtering model refines the recommendation allowing the model to suggest the most appropriate songs to the driver based on their mood.	-

to enhance the system's performance.

A. Ferraro *et al.* [10] investigates the impact of artist and style exposure bias on collaborative filtering-based music recommendations. The authors analyze a large-scale dataset of music listening histories and show that users tend to listen to popular artists and genres, which can lead to exposure bias in collaborative filtering-based recommendations. The authors have thought about how the popularity bias affects Matrix Factorization-based collaborative filtering recommendations and they have proposed an algorithm which is boosting the exposure of more well-known musical genres while decreasing the exposure in the long tail.

J. Singh [28] has described a hybrid music recommendation system with the goal of determining the likelihood that a user will listen to the song repeatedly after the first apparent listening experience begins in the time frame. This system uses both collaborative and content-based filtering algorithms. Additionally, methods including collaborative filtering, singular value decomposition, and factorization machines (FM) are employed. In order to increase the overall accuracy of the recommendation system utilising deep neural networks, the author has also hybridised FM and SVD models.

Joshi *et al.* [16] use Long Short-Term Memory (LSTM), Convolution Neural network (CNN), and LSTM-CNN architecture in a music recommendation system that uses deep learning techniques to detect emotions from facial expressions and recommend music that matches the detected emotion. The paper demonstrates the potential of using deep learning techniques for emotion-based music recommendation and highlights the importance of considering contextual information, such as facial expressions, in recommendation systems[6].

In [19] Xiaoyu Lin presents a comprehensive music recommendation system that utilizes Collaborative Filtering (User CF and Item CF) and Singular Value Decomposition (SVD) algorithms. Using the Jaccard index, the Item CF algorithm calculates the similarity between a user's songs and all unique songs. The User CF algorithm measures the similarity between users based on their song preferences. The SVD algorithm decomposes rating matrices to extract latent factors for personalized recommendations. Using the SVD and User CF to predict certain users tastes in music the accuracy is 80% and 88%.

3. Proposed Research Work. We propose our system in detail in the following subsections. These subsections would include a system architecture, a brief methodology of each individual component and finally, the amalgamation of these components into a unified framework for music song recommendation.

3.1. Dataset Description. The following section contains detailed description on the datasets chosen in accordance to the scope of the proposed method and tweaked to fit specfic requirements aimed by us during the inception of the methodology. We have selected the Facial Emotion Recognition dataset to learn an emotion-detecting model, and Spotify's Million Playlist Dataset to implement collaborative and content-based filtering models. Following subsections refer to them in detail.

3.1.1. Facial Emotion Dataset. The Facial Emotion Recognition - 2013 : (FER-2013) dataset [2] was chosen for our work, keeping in mind the requirements posed in detecting drivers' emotions and the effectiveness in capturing the said features. A subset of this dataset was taken that would better augment the application of detecting driver's emotions. Out of the seven put forward in the dataset, we selected four emotions that aligned with our proposed methodology; viz. happy, sad, angry and neutral. These emotions were identified as useful in detecting the mental state of the driver, which would then be useful in subsequently suggesting relevant songs to the driver. Some of the photos of the dataset are shown in Figure 3.1.

Our subset of the dataset consists of a total of 21,004 images spanning four categories, which are split into training and test datasets. The training dataset consists of 14,702 images and the test dataset consists of 6,302 images. The images present in the dataset are gray-scaled images of dimensions 48 x 48 pixels each. The gray-scaled images help reduce the learning load on the model by eliminating irrelevant data; eg. color, from the input. These images contain human faces that are mostly centred in the image, which helps in the identification of facial emotions easier as compared to faces that contain facial features at the corners or at some other position in the image.

3.1.2. Spotify's Million Playlist Dataset. The recommendation engine in this study utilizes the Spotify Million Playlist Dataset (MPD) to generate user profiles. This dataset encompasses one million playlists with over two million distinct tracks from nearly 300,000 artists. Each playlist includes details such as the playlist title, track list with corresponding track IDs and metadata, as well as additional metadata fields like last edit

Fig. 3.1: Some images from the Facial Emotion Recognition Dataset

time and the number of playlist edits. In this context, each playlist is treated as a user, and the songs within the playlist are considered to be the preferences of that hypothetical user. To simulate user preferences, certain songs have been augmented to the list of liked songs for these imaginary users, while other values have been assigned to indicate user dislike for specific songs. Liked songs are denoted as 1, while disliked songs are represented as -1. The resulting matrix is sparse, with the majority of values being 0, which signifies songs the user has not listened to or, in essence, potential recommendations. To facilitate efficient processing, a random subset of playlists and songs has been extracted from the dataset for our experimental analysis. This subset remains representative of the full dataset, capturing its crucial attributes for our analysis without compromising efficiency.

3.2. System Architecture. ConCollA music recommender system incorporates a multi-faceted approach to enhance the accuracy and relevance of song recommendations based on the driver's mood. The system architecture, as depicted in Figure 3.2 consists of several components, including mood recognition using a Convolutional Neural Network (CNN) model, collaborative-based recommendations, content-based recommendations, and association rule mining.

The initial step in the system is mood recognition, which relies on an in-car camera capturing images of the driver. These images are then fed into a CNN model trained to recognize patterns in photos corresponding to different emotions. The CNN model employed in this architecture consists of four convolutional layers and two dense layers. The system determines the driver's mood through the model, which serves as a crucial input for generating personalized song recommendations. Following that, the system proceeds to provide recommendations through a novel approach using a combination of collaborative filtering, content-based filtering along with associative rule mining. The collaborative-based recommendation component employs matrix factorization techniques to uncover underlying emotional factors that influence music preferences. By representing users and music items in emotional dimensions, the system generates personalized recommendations aligned with the user's emotional state and inclinations. This approach utilizes an anticipated rating matrix to deliver tailored suggestions based on the user's emotional profile and the content of the music.

Within the system architecture, we employ a content-based recommendation approach that leverages the acoustic features of music. Content-based recommenders analyze the intrinsic characteristics of songs, such as melody, rhythm, and instrumentation, to identify similarities between different tracks. In our case, we have developed a novel method called the Mood-Adjusted Average Similarity (MAAS) Algorithm.

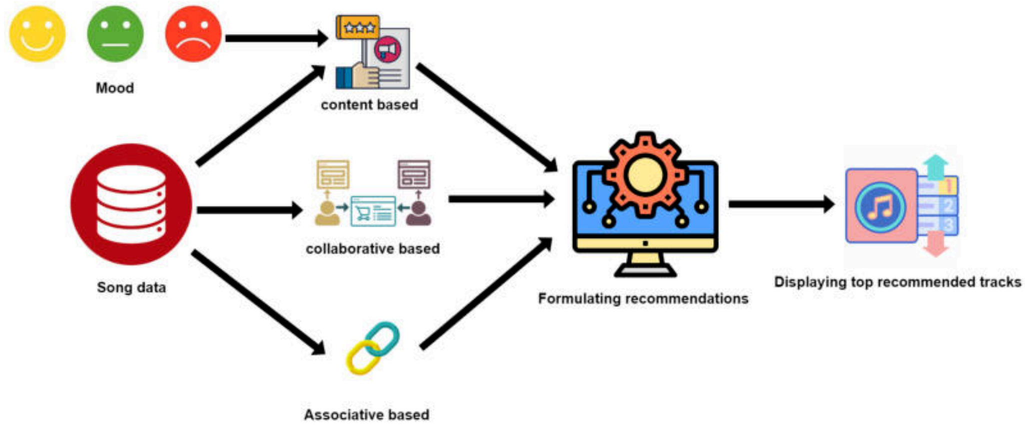

Fig. 3.2: System Architecture

The MAAS Algorithm takes into account the user's listening history and their emotional state. For each song that the user has not yet heard, we calculate its similarity to the songs they have previously listened to. This is achieved by computing the average similarity score between the target song and the songs heard by the user using acoustic feature analysis. The MAAS Algorithm goes a step further by involving mood biases into the picture, ensuring that the similarity score is adjusted according to the user's current emotional state. The resulting similarity scores range between -1 and 1, with higher values indicating a stronger match between the target song and the user's preferences. By considering both acoustic features and mood biases, the MAAS Algorithm enhances the accuracy and relevance of song recommendations, tailored to each individual driver's tastes and emotional context.

ConCollA also harnesses the power of association rule mining to further enhance the music recommendation process. While collaborative filtering focuses on finding recommendations from similar users, association rule mining allows us to uncover strong associations among songs on a larger scale, considering the listening patterns of a wide group of users. By applying association rule mining techniques, we can identify meaningful associations between songs that may not have been apparent through traditional collaborative filtering methods. These associations provide valuable insights into the preferences and listening behaviors of users. For example, if a particular song is frequently listened with another song, the association rule mining approach can identify this strong association. Consequently, the algorithm can deduce the presence of the consequent (the new song not yet heard) based on the presence of the antecedent (the song heard by the user), helping the former get additional preference due to said correlation.

This broader perspective enables our recommendation system to capture hidden relationships and co-occurrences among songs, enhancing the accuracy and relevance of the recommendations. By considering associations beyond the scope of individual user preferences, we can offer diverse and serendipitous recommendations that may align with the user's musical taste but also introduce new and exciting tracks.

By aggregating the results from these recommendation systems, the system generates a final recommendation rating. Based on the rating of each song, the system plays the recommended tracks to the driver in descending order, enhancing their music-listening experience during their journey.

The modular design and integration of these recommendation approaches within the system architecture ensure a comprehensive and personalized music recommendation process, taking into account user emotions, preferences, and associations to provide a tailored and enjoyable music experience.

4. Proposed Methodology. We introduce our proposed methodology for creating a novel recommender system that significantly improves the recommendation process by incorporating driver mood recognition in this section. The methodology comprises two key components: Driver Mood Recognition and Recommender Engine. The primary objective of the first component is to accurately capture the emotional state of the

Table 4.1: Layer information for the Emotion recognition model

Layers	**Output Shape**	**Parameters**
InputLayer	48,48,3	0
Conv2D	48,48,64	1792
BatchNormalization	48,48,64	256
Conv2D	48,48,64	36928
BatchNormalization	48,48,64	256
MaxPooling2D	24,24,64	0
Dropout	24,24,64	0
Conv2D	24,24,128	73856
BatchNormalization	24,24,128	512
Conv2D	24,24,128	147584
BatchNormalization	24,24,128	512
Conv2D	24,24,128	147584
MaxPooling2D	12,12,128	0
Dropout	12,12,128	0
Conv2D	12,12,256	295168
BatchNormalization	12,12,256	1024
Conv2D	12,12,256	590080
BatchNormalization	12,12,256	1024
Conv2D	12,12,256	590080
BatchNormalization	12,12,256	1024
Conv2D	12,12,256	590080
BatchNormalization	12,12,256	1024
MaxPooling2D	6,6,256	0
Dropout	6,6,256	0
Flatten	9216,1	0
Dense	4,1	36868

driver through the implementation of diverse techniques and sensors. Subsequently, this invaluable emotional information will be seamlessly integrated into the recommender engine, enabling the generation of personalized recommendations that precisely align with the driver's mood and preferences.

The integration of driver mood recognition with the recommender engine is crucial for creating a more engaging and satisfying user experience. By understanding the driver's mood, the recommender system can adapt its recommendations to match the driver's emotional state, thereby increasing the likelihood of a positive response and user satisfaction.

4.1. Driver Mood Recognition. The driver mood detection algorithm is a crucial element of the recommendation engine. It provides key information to the proposed recommender engine by infusing the knowledge of the driver's mood into the recommender, thereby better augmenting the data being used by the system. Providing this crucial insight into the moods of the driver has a significant importance on the selection of songs for the driver; improving the mental conditions and allowing the driver to focus on the roads. The architecture and description of the model follows.

4.1.1. Architecture. The mood recognition model is a convolutional model that is split into three blocks. Table 4.1 shows a detailed view of the model where all the added layers along with respective output shape and number of parameter is mentioned. The visualization of the network for emotion classification is depicted in Figure 4.1.

The input layer (in green) takes an input image of dimensions $48 \times 48 \times 3$. The first block has two convolutional layers (in orange) with 64 filters in each layer. These features are normalized using a Batch Normalization(BN) layer (shown in blue), followed by a MaxPooling Layer (in purple). Their resultant output is then sent through a dropout filter (in red) which provides a cap on the extent of learning, thus allowing the

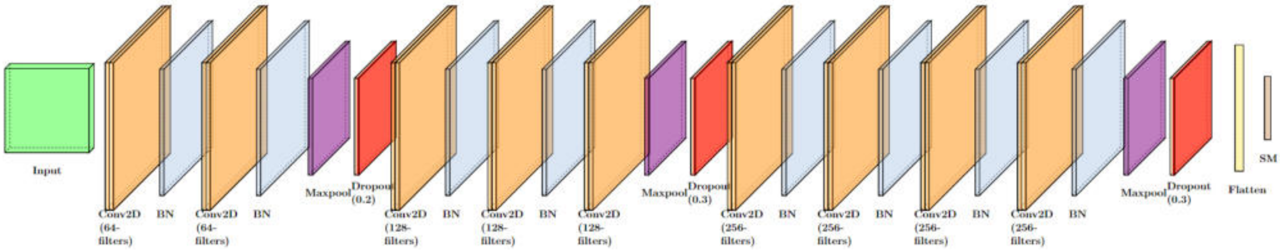

Fig. 4.1: CNN Architecture for Emotion recognition

model to better generalize on the training dataset.

The second block has the same architecture as the first block, namely the convolutional layer; but with 128 filters, a Batch Normalization layer and then a dropout layer. The rate of dropout is increased here to ensure only certain features from the previous block are passed on to the next block.

The final block has three convolutional layers of 256 filters each, along with a Batch Normalization layer after each convolutional layer and a dropout layer at the end. The dropout layer is then sent into a Flatten layer (in yellow) and fed into a fully-connected layer (in brown) containing four neurons as the final output. These neurons are activated by the softmax (SM) activation function that gives a likelihood score to each class.

4.2. Recommender Engine. The recommender engine is a pivotal component of ConCollA recommender system, responsible for generating personalized recommendations based on the driver's mood and preferences. This subsection outlines the methodologies and techniques employed within the recommender engine, including collaborative filtering, content-based recommender, and association rule mining.

4.2.1. Collaborative Filtering. Collaborative filtering is a noteworthy approach in recommendation systems for generating personalized recommendations by leveraging users' past behavior or preferences. This approach operates under the assumption that individuals with similar preferences in the past are likely to make similar choices in the future. Consequently, collaborative filtering recommends items that are popular among users with comparable preferences. There are many variants of collaborative filtering such as user-based, item-based, and model-based.

User-based collaborative filtering selects similar users based on past behavior or preferences and recommends items popular among them but not interacted with by the current user, aligning with collective preferences. Item-based collaborative filtering identifies similar items based on previous user interactions and suggests items similar to those already interacted with, leveraging item relationships to recommend related items.

Model-based collaborative filtering, such as Matrix Factorization, is another approach within collaborative filtering. This technique employs a mathematical model to uncover latent factors that capture the underlying user-item interactions. Matrix Factorization decomposes the user-item interaction matrix into lower-dimensional latent factor matrices. By utilizing these latent factors, personalized recommendations can be generated. Matrix Factorization aims to capture the hidden patterns and preferences within the user-item interactions, allowing for more accurate and effective recommendations.

Matrix factorization is a popular collaborative filtering method used to identify latent variables within user-item interactions. In an emotion-based music recommendation system, matrix factorization can uncover emotional content in music, enabling personalized recommendations based on the user's emotions and preferences. The Spotify Million Playlist Dataset provides this user-item matrix in the form of a user-track matrix.

The process involves breaking down the user-item interactions matrix into two lower-rank matrices, representing latent attributes of users and items. The user matrix indicates preferences across latent factors, while the item or track matrix represents the features of each item. By multiplying these matrices, a projected rating matrix is obtained, forming the basis for generating user-specific recommendations [29].

We utilized Singular Value Decomposition (SVD) [Algorithm 1], a widely adopted matrix factorization technique in collaborative filtering for ConCollA. SVD was utilized to decompose the user-item interaction matrix into three constituent matrices: U, Σ, and V. The user matrix U captures the preferences of each user, while the diagonal matrix Σ reflects the significance of the singular values obtained during the decomposition. Additionally, the item matrix V encapsulates the characteristics of each item. By leveraging the insights provided by the singular values derived from SVD, we identified crucial latent factors that play a significant role in explaining user-item interactions and generating accurate recommendations[20].

The process of calculating the matrices U, Σ, and V in SVD involves finding the eigenvectors and eigenvalues of the matrix M. The eigenvectors form the columns of U, while the eigenvalues form the diagonal elements of Σ. The columns of V are derived from the eigenvectors of the transposed matrix of M.

$$M = U\Sigma V^T \tag{4.1}$$

where:
- M is the user-item interaction matrix,
- U is the matrix of left singular vectors (representing users),
- Σ is the diagonal matrix of singular values,
- V is the matrix of right singular vectors (representing items), and
- V^T denotes the transpose of the matrix V.

The diagonal matrix Σ consists of singular values $(\sigma_1, \sigma_2, \ldots, \sigma_k)$ along its diagonal, where k is the rank of the matrix.

To calculate the matrices U and V, one can use eigenvalue decomposition or singular value decomposition algorithms. These algorithms provide the eigenvalues and eigenvectors, which can be used to construct the matrices U and V as follows:

$$U = \begin{bmatrix} u_1 & u_2 & \ldots & u_m \end{bmatrix} \tag{4.2}$$

$$V = \begin{bmatrix} v_1 & v_2 & \ldots & v_n \end{bmatrix} \tag{4.3}$$

where u_i and v_i represent the eigenvectors associated with the eigenvalues σ_i of Σ.

Using these matrices, the projected rating matrix $\hat{M}$ can be obtained as:

$$\hat{M} = U\Sigma V^T \tag{4.4}$$

The utilization of SVD in ConCollA enabled us to effectively extract latent factors from the user-item interaction matrix and leverage them to generate personalized recommendations. Its application in our recommender is as explained below.

Algorithm 1 Singular Value Decomposition (SVD)

Require: $m \times n$ matrix M

Ensure: Orthogonal matrices U and V, and diagonal matrix Σ such that $M = U\Sigma V^T$

1: Calculate $M^T M$, which is an $n \times n$ matrix.
2: Find the eigenvalues and eigenvectors of $M^T M$.
3: Sort the eigenvalues in descending order and form a diagonal matrix Σ, where the diagonal elements are the square roots of the eigenvalues in descending order.
4: Normalize the eigenvectors from the previous step to form an $n \times n$ orthogonal matrix V, where the columns are the normalized eigenvectors.
5: Calculate MV, which gives an $m \times n$ matrix.
6: Normalize each column of the matrix obtained in the previous step to obtain an $m \times n$ orthogonal matrix U.
7: **Return** the matrices U, Σ, and V^T.

Table 4.2: Musical Features

ID	Acousticness	Danceability	Energy	Key	Liveness	Loudness	Mode	Speechiness	Tempo
1	0.75	0.65	0.82	G	0.60	-5.2	Major	0.32	120
2	0.32	0.78	0.90	C	0.45	-3.8	Minor	0.15	130
3	0.91	0.40	0.67	A	0.70	-8.1	Major	0.22	105
4	0.60	0.85	0.76	E	0.55	-6.3	Minor	0.40	140

4.2.2. Content Based Recommender. The prevalent approach in modern music services is collaborative filtering, which generally yields recommendations of acceptable quality. However, this method exhibits limitations when faced with new users or emerging music, as the scarcity of user interaction information hampers its effectiveness. This scenario is referred to as the "cold start problem" in academic research [21]. To mitigate this challenge, the inclusion of similar sounding music becomes crucial. For instance, to acquaint listeners with lesser-known artists, recommendations may encompass the artist's works alongside similar sounding songs by more prominent artists.

Content-based recommenders present users with objects that exhibit similarity to items that have previously captured their interest[14]. Unlike collaborative filtering, which assesses object similarity based on user actions and similarities among users, content filtering evaluates similarity solely on the intrinsic characteristics of the objects themselves. Content filtering relies exclusively on objective attributes of the objects and remains independent of subjective user ratings.

Typically, these attributes are represented using a structured database table [24]. In the context of our study, we specifically require the inclusion of musical acoustic features. The acoustic features utilized in our analysis comprise acousticness, danceability, energy, instrumentalness, key, liveness, loudness, mode, speechiness, and tempo. Table 4.2 provides a subset of records showcasing a few musical features from the ones mentioned above, where each row corresponds to a distinct song and the columns denote the identifiers for the song as well as the aforementioned musical attributes. Within each record, a value is assigned to each attribute. To differentiate songs with identical names and facilitate the retrieval of associated attributes, a unique identifier known as ID is employed as a key. To generate these records, we utilize the capabilities of Spotify API. The user-track matrix used in Collaborative filtering can be used to fetch the song IDs of all the songs present in the playlist. These song IDs are provided further to Spotify API which reverts back with the acoustic features discussed above.

To facilitate the recommendation of songs, it is essential to establish a measure of similarity between these items. One commonly employed method, as elucidated by [24], is the nearest neighbor method. The choice of similarity function in the nearest neighbor algorithm depends on the data type under consideration. In the case of structured data, a frequently used metric is the Euclidean distance, whereas the vector space model often employs the cosine similarity measure.

The Euclidean distance function treats a small value for the same feature in two examples as equivalent to a large value for that feature in both examples. Conversely, the cosine similarity function does not assign a large value if corresponding features of two examples possess small values. Given that we treat the musical features as vectors in our task, the cosine similarity metric is employed. However, we propose a novel approach distinct from the native top-k similarity recommendation to leverage this cosine similarity and generate more effective recommendations.

In order to provide recommendations for unheard songs, it is necessary to partition the dataset into two distinct sets: Set P, consisting of songs the user has listened to, and Set Q, encompassing songs the user has not yet heard, for which predictions are required. To make recommendations, a crucial step involves calculating the similarity between each song in Set Q and every item in Set P. The recommendation score for a song q belonging to Set Q is then determined as the average of the similarities between q and each song p belonging to Set P.

This approach ensures that the recommendation score for an unheard song takes into account its similarity to the songs the user has already listened to. By calculating the similarity between q and each item in Set P, the recommendation score reflects the collective resemblance of q to the songs in the user's listening history. The average of these similarities provides an overall measure of recommendation, enabling the system to suggest songs in Set Q based on their similarity to the user's previously enjoyed songs.

Algorithm 2 Mood-Adjusted Average Similarity (MAAS)

Input: $P \leftarrow$ Set of Tracks User has listened to
$Q \leftarrow$ Set of Tracks to Recommend to User
$MoodCluster \leftarrow$ Set of Tracks corresponding to User's mood
Output: MAAS score

procedure SIMILARITY_CALCULATOR($Set\ P,\ Set\ Q$)
  for $q \in Q$ **do**
    $MAAS\_score[q] \leftarrow 0$ ▷ Initialize MAAS score for track q
    $q_{acc} \leftarrow$ get_acoustics(q) ▷ Extract acoustic features for track q
    for $p \in P$ **do**
      $p_{acc} \leftarrow$ get_acoustics(p) ▷ Extract acoustic features for track p
      $MAAS\_score[q] \leftarrow MAAS\_score[q] +$ cosine_similarity(p_{acc}, q_{acc}) ▷ Calculate cosine similarity
    end for
    $MAAS\_score[q] \leftarrow MAAS\_score[q]$/length($P$) ▷ Normalize MAAS score
    if $q \notin$ MoodCluster **then**
      $MAAS\_score[q] \leftarrow -MAAS\_score[q]$ ▷ Adjust MAAS score based on mood
    end if
  end for
end procedure

The rationale behind this methodology is to leverage the known preferences of the user, as represented by Set P, to estimate the user's potential interest in the unheard songs within Set Q. By considering the similarities with the familiar songs, the recommendation system can make informed predictions about the user's taste and provide relevant recommendations accordingly. The averaging of similarity scores allows for a comprehensive assessment of the compatibility between the unheard songs and the user's established preferences.

The previously mentioned similarity score, while significant, does not encompass the user's mood consideration. To incorporate the user's mood into the averaged recommendation score, an additional step is undertaken. This process involves iterating through each song and adjusting the recommendation score for songs that do not align with the user's mood. The score that we receive now is called MAAS score, as mentioned in the architecure section. This score is calculated through [Algorithm 2]. By negating the recommendation score, the resulting score falls within the range of -1 to 0, considering that the original recommendation score ranges from 0 to 1. When combining scores obtained from different methods, this negated score contributes to reducing the effective score for songs that do not align with the user's mood. This mechanism accounts for the user's mood preferences which contributes to a more personalized and mood-aligned recommendation experience.

4.2.3. Association Rule Mining. Association rule mining is a data mining technique used to uncover meaningful associations or patterns in large datasets. In the context of music recommendation systems, association rule mining aims to discover relationships between music items based on their co-occurrence patterns. By analyzing user listening behavior, it can identify frequently occurring combinations of songs or artist preferences. The goal is to extract association rules that capture the relationships between music items, allowing us to understand the preferences and associations of different users within the music domain. These rules can reveal interesting connections, such as songs that tend to be listened together or a group of people frequently listening to the same artist.

In a music recommendation system, association rule mining can improve the diversity and personalization of recommendations[15]. The system can make related or complementary song suggestions by taking into consideration these discovered relationships, enhancing the user's ability to find new music. Additionally, it offers personalised playlists catered to each listener's preferences, suggests music based on their likes, and aids in identifying up-and-coming artists.

Association rule mining's use in music recommendation systems creates new opportunities for offering interesting and relevant musical suggestions. The algorithm may generate more individualised and varied recommendations that are tailored to each user's own tastes and preferences by understanding the relationships between musical pieces.

Association rule mining techniques, such as the Apriori algorithm [Algorithm 3], are commonly employed in music recommendation systems to discover meaningful relationships among songs. The Apriori algorithm, firstly conceptualized by Agrawal et al., is a popular method for extracting frequent itemsets and deriving association rules.

The Apriori algorithm follows a level-wise search strategy to efficiently identify frequent itemsets. It starts with frequent itemsets of length 1, also known as "singleton" itemsets, and progressively generates candidate itemsets of increasing lengths. The algorithm prunes the search space by eliminating itemsets that cannot be frequent based on the infrequency of their subsets, leveraging the apriori property.

In the context of music recommendation, we use the Apriori algorithm to transactional data representing user playlists. This transactional data is provided by Spotify Million Playlist Dataset in form of user-track matrix, which is the same used by the Collaborative recommender. By identifying frequent itemsets, the algorithm uncovers associations between songs or artists. These associations help enhance the accuracy and diversity of music recommendations by suggesting related or complementary music items to users [30].

Algorithm 3 Apriori Algorithm

$Input \leftarrow$ Load User-Playlist dataset
$Output \leftarrow$ Large Itemsets
$L_1 =$ large 1-itemsets ▷ Initialize with 1-itemsets
for $k = 0; L_{k-1} \neq \emptyset; k++$ **do** ▷ Iterate until no more frequent itemsets
    $C_k = \text{apriori\_gen}(L_{k-1})$ ▷ Generate candidate itemsets
    for $\forall$transaction $t \in D$ **do** ▷ Scan transactions
        $C_t = \text{subset}(C_k, t)$ ▷ Find candidate subsets in transaction
        for candidate $c \in C_t$ **do** ▷ Count occurrences of candidates
            c.count++
        end for
        $L_k = \{c \in C_k \mid c.\text{count} \geq \text{min\_sup}\}$ ▷ Keep frequent itemsets
    end for
end for
Answer $= U_k L_k$ ▷ Final set of large itemsets

Once the frequent itemsets have been identified using the Apriori algorithm, the next step is to derive association rules from these itemsets. Association rules capture the relationships and dependencies among items based on their co-occurrence patterns. In the context of ConCollA, these association rules provide valuable insights into the preferences and associations within the music domain. Using association rule mining, it is possible to find patterns in massive amounts of data. In order to enable analysts to meaningfully connect seemingly unrelated pieces of data, it helps to discover co-occurring events or objects that might not be immediately obvious. Here it helps us to find the combination of occurrences of various tracks together and gives us a relationship stating if these n tracks are in a playlist then these m tracks can also be recommended [12].

To generate association rules, the frequent itemsets serve as the basis. Each frequent itemset represents a set of items that occur together frequently in user playlists or listening histories. From these itemsets, association rules are derived by considering various combinations of the items within each frequent itemset.

The process of generating association rules involves defining thresholds or measures of interestingness, such as support, confidence, and lift. Support measures the frequency of occurrence of an itemset in the dataset, confidence quantifies the strength of the rule by measuring the conditional probability, and lift determines the degree of association between items [18].

The culmination of the associative rule mining phase marks the completion of all individual components within the recommendation system. The upcoming crucial step involves the integration of results derived from three distinct techniques: content-based filtering, collaborative filtering, and association rule mining. The combined implementation of these three approaches facilitates the recommendation process, enabling the system to offer personalized music recommendations based on users' moods. The integration of these methodologies ensures a comprehensive and refined approach to music recommendation. By considering users' preferences, past behavior, and the discovered associations among music items, ConCollA aims to deliver accurate and personalized music suggestions. In the subsequent section, we will discuss upon the details of this integration and examine its impact on the recommendation engine. This analysis will reveal how the ConCollA recommender system intelligently leverages content-based, collaborative, and association rule mining techniques, resulting in a highly effective and user-centric music recommendation system.

4.3. Integration. In order to integrate the individual components of ConCollA, viz. collaborative filtering, content-based filtering, and association rule mining, we employ a weighted aggregation approach that combines scores from collaborative and content-based filtering techniques by aggregating the scores after normalizing them with the respective weights. The following formula represents the weighted aggregation of content-based and collaborative filtering techniques:

$$\mathrm{RS} = \mathrm{W_{CB}} \times \mathrm{Score_{CB}} + \mathrm{W_{CO}} \times \mathrm{Score_{CO}} \tag{4.5}$$

where RS is the normalized recommendation score, $\mathrm{W_{CB}}$ is the weight assigned to content-based filtering model, $\mathrm{Score_{CB}}$ is the score given by the content-based model, $\mathrm{W_{CO}}$ is the weight assigned to collaborative filtering model, $\mathrm{Score_{CO}}$ is the score given by the collaborative model.

In ConCollA recommender system, association rule mining is incorporated to further enhance the recommendation process. By analyzing the user's listening history, ConCollA checks the presence of antecedent items (X) from association rules. If the antecedent is found, we add the normalized confidence score of the association rule to the recommendation score of the consequent items (Y). This allows ConCollA to leverage the strength of the association rules while considering the user's preferences. To ensure a balanced influence, ConCollA normalizes the confidence score within the range of 0 to 0.1, preventing it from dominating the overall recommendation results. The combination of content-based filtering, collaborative filtering, and association rule mining enables ConCollA to provide accurate and personalized music recommendations to users. So if a rule holds true, the above given Recommendation Score formula is changed as follows:

$$\mathrm{RS} = \mathrm{W_{CB}} \times \mathrm{Score_{CB}} + \mathrm{W_{CO}} \times \mathrm{Score_{CO}} + \mathrm{W_{A}} \times \mathrm{Score_{ARC}} \tag{4.6}$$

where $\mathrm{W_A}$ is the weight factored to normalize the associative score, $\mathrm{Score_{ARC}}$ is the associative rule confidence for the antecedent.

The weights assigned to different recommendation components play a crucial role in providing personalized recommendations to users. In ConCollA, we recognize the diversity of user preferences and configure individual weights for each user to tailor the recommendations accordingly. For example, while collaborative filtering identifies similar users, we acknowledge that some users may have a preference for songs similar to their own, which are provided by the content-based recommendation component. Hence, such users receive higher weights on the content-based recommendation, emphasizing the importance of recommendations based on content similarity. Conversely, users who rely more on collaborative filtering receive higher weights for that component.

The determination of these weights can be framed as a linear regression problem. We aim to find the optimal weights that minimize the difference between the predicted recommendation scores and the user's actual preferences. This approach allows us to model the relationship between the recommendation components and the user's feedback, facilitating the fine-tuning of the weights to align with their preferences.

To optimize the weights for each user, ConCollA employs the gradient descent algorithm. Gradient descent is an iterative optimization technique that adjusts the weights based on the calculated gradients of the loss function. By updating the weights in the direction that minimizes the loss in each iteration, we can converge towards

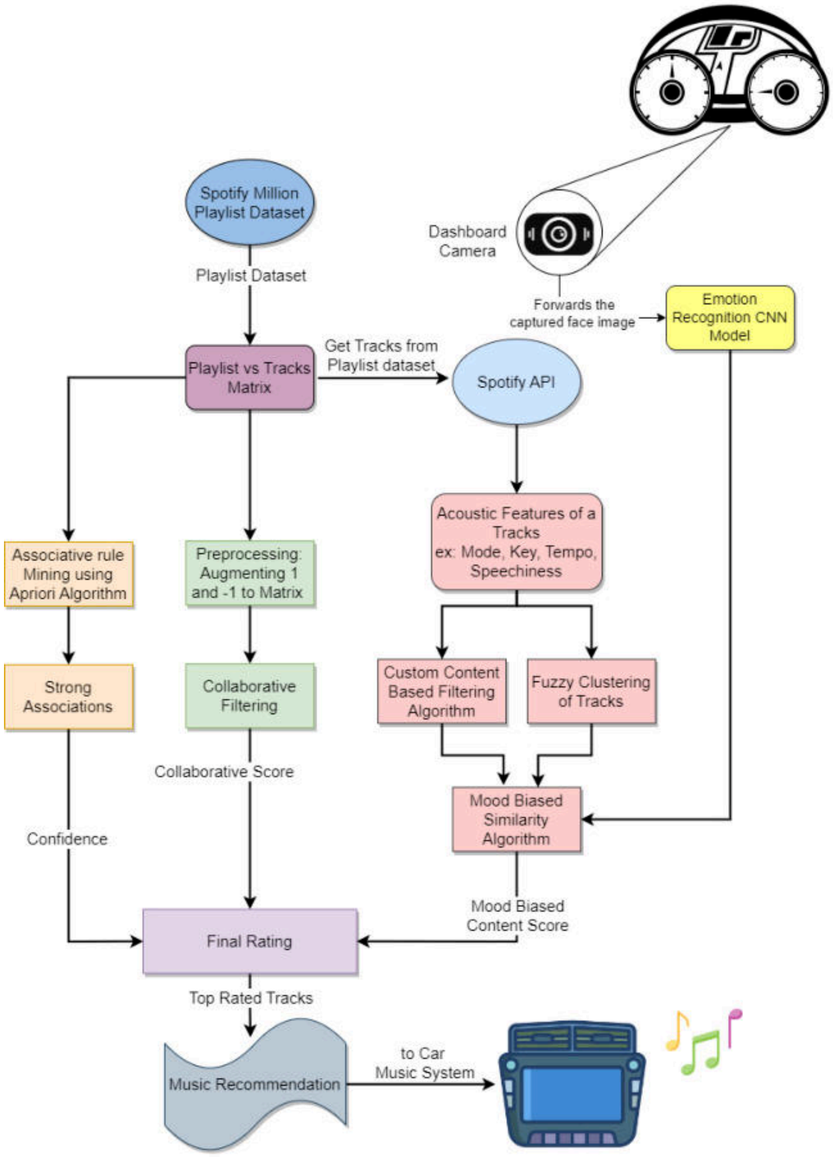

Fig. 4.2: Flow diagram of the model

the optimal set of weights for each user, maximizing the accuracy and effectiveness of the recommendations. The generic formula for gradient descent in our music recommendation context can be written as:

$$\theta_j := \theta_j - \alpha \frac{\partial J(\theta)}{\partial \theta_j} \tag{4.7}$$

In this formula, Weight_j represents the weight for a specific recommendation component, such as content-based, collaborative filtering, or association rules, that needs to be updated. α is the learning rate, controlling the step size of the weight update. $\frac{\partial J}{\partial \text{Weight}_j}$ denotes the partial derivative of the cost function J with respect to the weight Weight_j.

ConCollA aims to optimize these weights for each user to provide personalized recommendations. The weights are initially assigned arbitrary values and are gradually updated using the gradient descent algorithm. This iterative optimization process allows us to adjust the relative importance of each recommendation component based on individual user preferences and feedback.

To optimize the weights, ConCollA defines a cost function that quantifies the discrepancy between the predicted recommendation scores and the user's actual preferences. This cost function serves as a measure of the accuracy or quality of the recommendations. The gradient descent algorithm calculates the gradients of the cost function with respect to each weight and updates the weights in the direction that minimizes the cost function.

The mean squared error (MSE) is the cost function that is most frequently employed in linear regression problems. The MSE calculates the average squared difference between the training data's actual values and the values that were predicted. It measures the linear regression model's overall goodness of fit.

By updating the weights using gradient descent after each iteration, we refine the recommendation process to better align with the user's preferences. The learning rate α controls the step size of the weight update, ensuring a balance between convergence speed and stability. ConCollA continues iterating until the weights converge to a point where the recommendation scores align closely with the user's preferences.

It is important to note that the weights are not fixed and are constantly updated as users interact with the recommended songs. This adaptive approach ensures that the recommendations remain aligned with the user's evolving preferences and feedback. By continuously optimizing the weights through gradient descent, our system adapts to changes in user preferences, providing more accurate and personalized recommendations over time. The aforementioned methodology has been explained in Fig. 4.2. Our proposed system can be used to generate music recommendations for drivers. It monitors the driver's emotions using a dashboard camera positioned in front of the driver and the collected emotional data from the camera is combined with past listening preferences to generate music recommendations. The suggested song is played automatically in the car and improves the listener experience.

5. Experimental Analysis. This section is divided into two subsections. The first subsection discusses the Evaluation Measures used for evaluating the proposed recommender system. The subsequent section then discusses the results achieved on the evaluation measures discussed in the former subsection.

5.1. Evaluation Parameters. For evaluation of the recommender, two metrics are provided: Root Mean Square Error (RMSE) and a novel metric, Rel-Sim.

Root Mean Square Error (RMSE) calculates the square root of the average of the squared differences between the predicted values and the actual (observed) values. In other words, it measures how far, on average, the predictions are from the actual values. A lower RMSE indicates better model performance, as it signifies that the predictions are closer to the true values [27].

$$\text{RMSE}(y, \hat{y}) = \sqrt{\frac{\sum_{i=0}^{N-1}(y_i - \hat{y}_i)^2}{N}} \tag{5.1}$$

Rel-Sim, which stands for "Relative Similarity" is a novel metric proposed for evaluating the ConCollA system, focusing on the relationships between the ground truth and the top recommended items. While conventional evaluation metrics like RMSE may yield low scores for the ground truth, they might not reflect its true relevance due to the influence of large datasets, resulting in lower MAP scores. To address this challenge, ConCollA evaluation metrics introduce Rel-Sim, which measures the similarity between the top-N recommended items and the ground truth using acoustic cosine similarity. Each item in the top-N recommendations, is compared with the ground truth and its acoustic cosine similarity is calculated. We then identify the ground truth item with the maximum similarity and multiply this value by the predicted score of the ground truth. This process is repeated for all top-N items, and the resulting values are averaged to obtain the final Rel-Sim score. A higher Rel-Sim score signifies that the top-N recommended items are highly similar to the ground truth, even if the ground truth item itself was not included in the top-N recommendations. Overall, Rel-Sim provides valuable insights into the recommender system's performance, particularly when the ground truth items are not among the top recommendations.

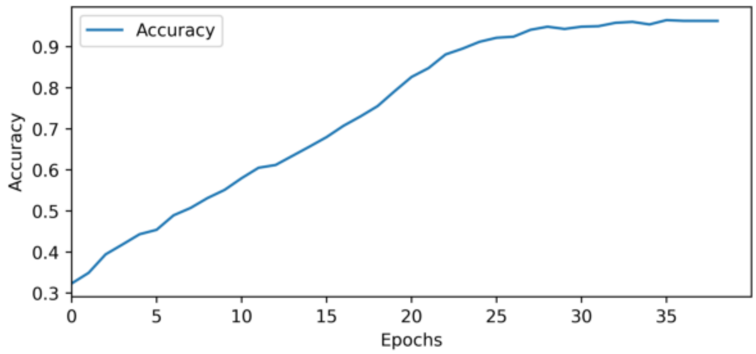

Fig. 5.1: Evaluation metrics for emotion recognition

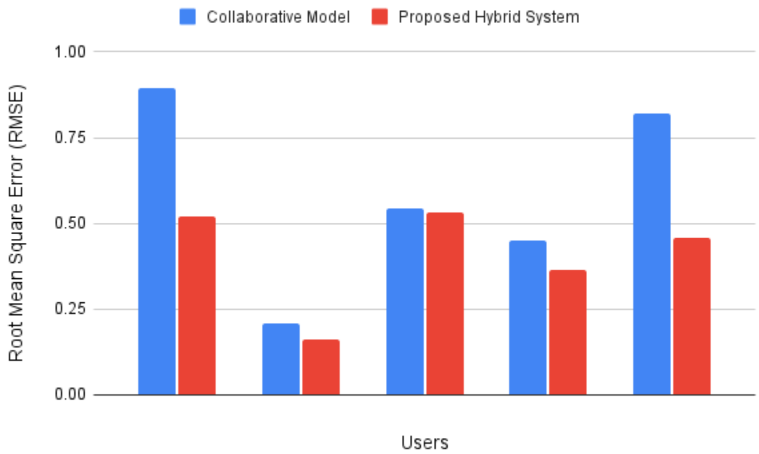

Fig. 5.2: RMSE (Lower is Better)

5.2. Results and Discussion. The ConCollA system is going to be evaluated under two phases, i.e. mood recognition and mood-based music recommendation. In the first phase, the emotion of the driver is detected and in the second phase, music is recommended based on the detected emotion. The results of each phase after evaluating them are discussed below.

5.2.1. Phase-1. The metric used for evaluating the emotion recognition model is the training accuracy. We have taken a subset of the FER-dataset based on our application. We run the model for 40 epochs and the results are obtained are for the same number of epochs. A plot of the graph of accuracy vs. number of epochs is shown in Fig. 5.1.

We have achieved training accuracy of 96.22% on the FER-2013 dataset over the duration of 40 epochs.

5.2.2. Phase-2. In this phase, we delve into the metrics associated with the mood-based recommender engine, analyzing and evaluating its performance.

Collaborative filtering is one of the most common recommenders used everywhere. The below mentioned Table5.1 contains the results obtained from the collaborative recommender for 5 users. These users were picked randomly from a set of users. For each user two metrics are provided: Root Mean Square Error (RMSE)and a novel metric, Rel-Sim.

It is observed that users who have provided a larger amount of data in the form of likes and dislikes of songs

Table 5.1: Comparative Results for Recommender Systems

Users	Collaborative Filtering		Weighted Content+Collaborative Filtering	
	RMSE	RelSim	RMSE	RelSim
1	0.8938	0.5691	0.5184	0.7482
2	0.2097	0.7841	0.1625	0.8639
3	0.5447	0.8108	0.5333	0.8245
4	0.4488	0.8598	0.3626	0.8914
5	0.8219	0.5774	0.4588	0.7708

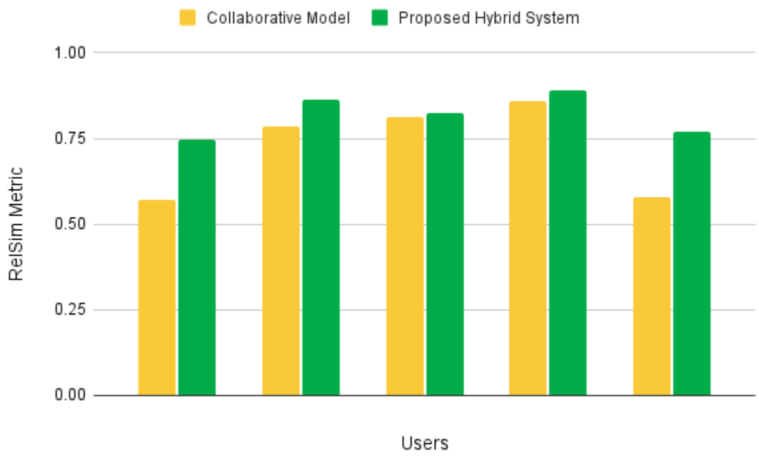

Fig. 5.3: Rel-Sim (Higher is Better)

tend to have lower RMSE and better Rel-Sim scores. However, there are a few users which have not provided sufficient data for recommendation,are not receiving recommendations as effectively as others, indicating a lower performance in terms of RMSE and Rel-Sim. This discrepancy can be attributed to the ”cold start problem”, where users who have not rated a sufficient number of songs pose a challenge for collaborative filtering algorithms to generate accurate recommendations.

To address the aforementioned problem,and include a mood bias in song recommendation, the proposed work incorporates a novel recommender system to augment the collaborative model. The Table 5.1 also presents the evaluation metrics for the proposed recommender system. The model observes an improvement in both the RMSE and Rel-Sim values for the majority of tested users. This improvement can be attributed to the personalized weightings assigned to each recommendation type. For instance, users X and Y have higher weights assigned to the content-based recommendations, indicating that they will receive a greater emphasis on content-based recommendations for songs. By fine-tuning these weights according to the user's preferences, we can enhance the accuracy and effectiveness of the recommendations provided.

Figure 5.2 displays the RMSE values in a graphical format, facilitating the interpretation of data presented in Table 5.1. The results demonstrate the superior performance of our proposed model, which exhibits lower RMSE values across all 5 randomly chosen users. Particularly noteworthy improvements are observed for Users 1 and 5, showcasing the efficacy of our approach. Similarly, Figure 5.3 presents the RelSim values for the aforementioned users, corresponding to the data in the respective table. Remarkably, the RelSim values have significantly increased for all users, providing further empirical support for our underlying hypothesis.

In addition to the hybrid recommender systems, ConCollA also incorporates strong association rules to provide a broader perspective on musical recommendations. These association rules capture patterns and

Table 5.2: Effect of adding Association Rule Mining on RMSE

	RMSE	
Users	Without Rule Mining	With Rule Mining
1	0.5184	0.5053
2	0.1625	0.1625
3	0.5333	0.5333
4	0.3626	0.3626
5	0.4588	0.4588

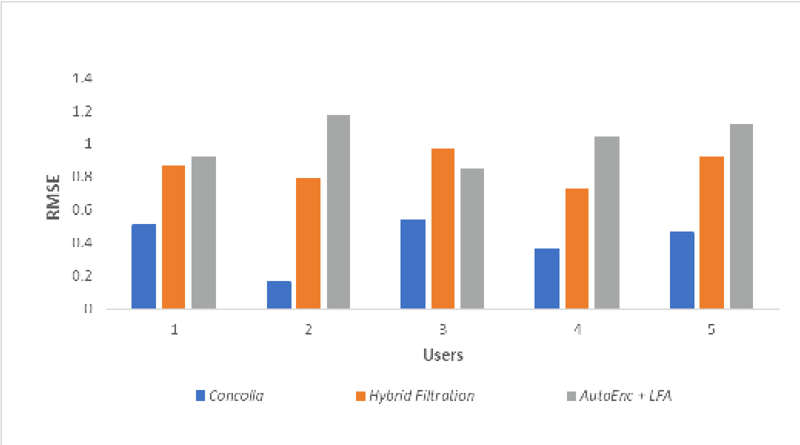

Fig. 5.4: RMSE (Lower is Better)

relationships among users beyond those who have similar preferences to the user being tested. Table 5.2 shows the comparisions between RMSE values of our approach without using strong associations while recommending and RMSE values for the users after adding a component of strong association rules to the former recommender approach. It can be noticed that User 1's RMSE value have decreased, whilst the other users have not changed. This is because there was a strong rule in the ground truth of 1 that led to a fall in its RMSE, while the users who were unaffected by it were because there was no strong rule.

We compare the performance of our model with two other models as depicted in Table 5.3. The first method is based on the concept of hybridization of a collaborative filtering recommender system and a music gene-based recommendation model [32]. For music gene-based recommendation, this approach considered characteristics of music such as tempo, rhythm, and various acoustic features and performed clustering on these characteristics. The results provided collaborative filtering and gene-based recommendation model are combined and filtered to generate one recommendation. Recommendation results can further be improved by finding the correlation among features. Hence we use association rule mining in our proposed model to find the correlation among features and therefore, our proposed model provides better personalized recommendations to the user. Along with that, we also propose our custom content-based algorithm, MAAS, which considers user's current emotional data while making recommendations. Figure 5.4 shows the enhanced outcomes of our proposed model ConColla by applying association rule mining through RMSE value.

Another method utilizes a hybrid approach termed AutoLFA, which aggregates two approaches, AutoEncoder and Latent Feature Analysis(LFA) to formulate a recommendation score. It aggregates these scores based on a customized self-adaptive weighting strategy [11]. We have also utilized this concept of sefl-adaptive weights in our model while combining our three individual systems, viz. collaborative recommendations, content-based recommendations and association rule mining. In our proposed approach, the scores from the three recommenders are aggregated in a weighted manner, and these weights are learnt differently for each user and are

Table 5.3: Comparison of ConCollA with other models

	RMSE		
Users	ConCollA	Hybrid Filtration	AutoEncoder + LFA
1	0.5053	0.867	0.925
2	0.1625	0.798	1.174
3	0.5333	0.975	0.852
4	0.3626	0.732	1.051
5	0.4588	0.921	1.128

self-adaptive and constantly changing as the user interacts with the system. To mitigate the potential cold start issue that could arise within this algorithm, we incorporate a content-based recommendation alongside the collaborative filtering algorithm in our proposed approach.

We employed these models as baseline models for our proposed system. We evaluated these models by training them on the dataset we used for our approach. A side-by-side comparison of RMSE metric of the five users used in previous experiments is shown for our proposed method and the two baseline methods in Fig. 5.4 as well as Table 5.3. The plot clearly shows that our proposed approach ConCollA is superior to both its predecessors for all the five randomly chosen users. This shows the efficacy of our approach combining various techniques to provide relevant recommendations.

6. Conclusion and Future Work. Music is one of the factors for improving the driving experiences, in this paper authors proposed a recommender system to make driving more enjoyable for drivers. ConCollA is a novel recommender system proposed based on personalized content based filtering, collaborative matrix factorization filtering followed by apriori algorithm from association rule mining through a multifaceted system design.A key feature of the model is the novel measure identification - Mood Adjusted Average Similarity (MAAS), and apriori algorithm to generate personalized music recommendations The success of the proposed model has been proved by the various test using a dataset from the Spotify API. Both the RMSE and Rel-Sim measures show that system outperforms the collaborative model, This method has demonstrated the ability to produce a more pleasurable and personalised music experience for drivers by fusing emotional awareness with conventional music selection algorithms. Despite the sophistication of music recommendation algorithms, very few have mastered the emotional power of music whilst driving. The proposed music recommendation system for drivers will pave the way for future advancements in the field of personalized and emotion-aware technology. The research gaps in ConCollA involve limited emotion understanding, overlooking individual variations, neglecting contextual information, challenges in quantifying subjective emotions, complex evaluation metrics, the need for cross-cultural and multilingual adaptability, and addressing ethical implications. Filling these gaps can lead to more personalized and intelligent music recommendations, enhancing user experience and system efficacy. Future research aims to incorporate contextual elements like weather, traffic, and time of day to improve the recommendation system. Additionally, efforts will focus on integrating complex emotional states and classifying songs based on individual user preferences, leading to more contextually aware and personalized music suggestions.

REFERENCES

[1] *World report on road traffic injury prevention.* `https://www.who.int/publications/i/item/world- report-on-road-traffic-injury-prevention`, 2004.

[2] *Goodfellow, et. al.: Challenges in representation learning: Facial expression recognition challenge.* `https://www.kaggle.com/competitions/challenges-in-representation-learning-facial-expression-recognition-challenge/data`, 2023.

[3] A. Abdul, J. Chen, H.-Y. Liao, and S.-H. Chang, *An emotion-aware personalized music recommendation system using a convolutional neural networks approach*, Applied Sciences, 8 (2018), p. 1103.

[4] W. G. Assuncao, L. S. Piccolo, and L. A. Zaina, *Considering emotions and contextual factors in music recommendation: a systematic literature review*, Multimedia Tools and Applications, 81 (2022), pp. 8367–8407.

[5] L. Baltrunas, M. Kaminskas, B. Ludwig, O. Moling, F. Ricci, A. Aydin, K.-H. Lüke, and R. Schwaiger, *Incarmusic: Context-aware music recommendations in a car*, in E-Commerce and Web Technologies: 12th International Conference,

EC-Web 2011, Toulouse, France, August 30-September 1, 2011. Proceedings 12, Springer, 2011, pp. 89–100.

[6] P. Bharati and A. Pramanik, *Deep learning techniques—r-cnn to mask r-cnn: a survey*, Computational Intelligence in Pattern Recognition: Proceedings of CIPR 2019, (2020), pp. 657–668.

[7] E. Çano, R. Coppola, E. Gargiulo, M. Marengo, and M. Morisio, *Mood-based on-car music recommendations*, in Industrial Networks and Intelligent Systems: Second International Conference, INISCOM 2016, Leicester, UK, October 31–November 1, 2016, Proceedings 2, Springer, 2017, pp. 154–163.

[8] J. J. Deng and C. Leung, *Emotion-based music recommendation using audio features and user playlist*, in 2012 6th International Conference on New Trends in Information Science, Service Science and Data Mining (ISSDM2012), IEEE, 2012, pp. 796–801.

[9] N. Dibben and V. J. Williamson, *An exploratory survey of in-vehicle music listening*, Psychology of Music, 35 (2007), pp. 571–589.

[10] A. Ferraro, D. Bogdanov, X. Serra, and J. Yoon, *Artist and style exposure bias in collaborative filtering based music recommendations*, arXiv preprint arXiv:1911.04827, (2019).

[11] S. Guo, X. Liao, G. Li, K. Xian, Y. Li, and C. Liang, *A hybrid recommender system based on autoencoder and latent feature analysis*, Entropy, 25 (2023), p. 1062.

[12] M. Henry, W. Chandra, and A. Zahra, *Implementation of apriori algorithm for music genre recommendation*, Jurnal Online Informatika, 7 (2022), pp. 110–115.

[13] A. V. Iyer, V. Pasad, S. R. Sankhe, and K. Prajapati, *Emotion based mood enhancing music recommendation*, in 2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), IEEE, 2017, pp. 1573–1577.

[14] U. Javed, K. Shaukat, I. A. Hameed, F. Iqbal, T. M. Alam, and S. Luo, *A review of content-based and context-based recommendation systems*, International Journal of Emerging Technologies in Learning (iJET), 16 (2021), pp. 274–306.

[15] J. Jooa, S. Bangb, and G. Parka, *Implementation of a recommendation system using association rules and collaborative filtering*, Procedia Computer Science, 91 (2016), pp. 944–952.

[16] S. Joshi, T. Jain, and N. Nair, *Emotion based music recommendation system using lstm-cnn architecture*, in 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), IEEE, 2021, pp. 01–06.

[17] A. S. Krishnan, X. Hu, J.-q. Deng, L. Zhou, E. C.-H. Ngai, X. Li, V. C. Leung, and Y.-k. Kwok, *Towards in time music mood-mapping for drivers: A novel approach*, in Proceedings of the 5th ACM Symposium on Development and Analysis of Intelligent Vehicular Networks and Applications, 2015, pp. 59–66.

[18] T. A. Kumbhare and S. V. Chobe, *An overview of association rule mining algorithms*, International Journal of Computer Science and Information Technologies, 5 (2014), pp. 927–930.

[19] X. Lin, *Music recommendation system based on collaborative filtering and singular value decomposition*, in CAIBDA 2022; 2nd International Conference on Artificial Intelligence, Big Data and Algorithms, VDE, 2022, pp. 1–5.

[20] N. Nassar, A. Jafar, and Y. Rahhal, *A novel deep multi-criteria collaborative filtering model for recommendation system*, Knowledge-Based Systems, 187 (2020), p. 104811.

[21] A. Niyazov, E. Mikhailova, and O. Egorova, *Content-based music recommendation system*, in 2021 29th Conference of Open Innovations Association (FRUCT), IEEE, 2021, pp. 274–279.

[22] J. Patel and H. Chhinkaniwala, *A fusion of aspect and contextual information for rating prediction in recommender system using a latent factor model*, International Journal of Web Engineering and Technology, 16 (2021), pp. 30–52.

[23] ———, *The role of sentiment analysis in a recommender system: a systematic survey*, International Journal of Web Engineering and Technology, 17 (2022), pp. 29–62.

[24] M. J. Pazzani and D. Billsus, *Content-based recommendation systems*, in The adaptive web: methods and strategies of web personalization, Springer, 2007, pp. 325–341.

[25] F. Richter, *Emotion detection: Fer - 2013 dataset with 7 emotion types.* `https://www.statista.com/chart/18557/where-and-when-people- listen-to-music/`, 2019.

[26] T. Schäfer, P. Sedlmeier, C. Städtler, and D. Huron, *The psychological functions of music listening*, Frontiers in psychology, 4 (2013), p. 511.

[27] T. Silveira, M. Zhang, X. Lin, Y. Liu, and S. Ma, *How good your recommender system is? a survey on evaluations in recommendation*, International Journal of Machine Learning and Cybernetics, 10 (2019), pp. 813–831.

[28] J. Singh, *Collaborative filtering based hybrid music recommendation system*, in 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS), IEEE, 2020, pp. 186–190.

[29] J. Singh and M. Sajid, *Factorization machine based music recommendation approach*, in 2021 6th International Conference on Inventive Computation Technologies (ICICT), IEEE, 2021, pp. 618–622.

[30] P. K. Singh, E. Othman, R. Ahmed, A. Mahmood, H. Dhahri, and P. Choudhury, *Optimized recommendations by user profiling using apriori algorithm*, Applied Soft Computing, 106 (2021), p. 107272.

[31] M. D. Van Der Zwaag, C. Dijksterhuis, D. De Waard, B. L. Mulder, J. H. Westerink, and K. A. Brookhuis, *The influence of music on mood and performance while driving*, Ergonomics, 55 (2012), pp. 12–22.

[32] D. Wu, *Music personalized recommendation system based on hybrid filtration*, in 2019 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS), IEEE, 2019, pp. 430–433.

Edited by: Rajni Mohana
Special issue on: Sentiment Analysis and Affective computing in Multimedia Data on Social Network
Received: Jul 31, 2023
Accepted: Sep 26, 2023