Background Research

1. Literature Review :

In this background research we have used literature survey of Our Projects research paper author Hareesh Bahuleyan, “Music Genre Classification using Machine Learning Techniques ”. This Excerpt is directly from the author’s Research paper.

Music genre classification has been a widely studied area of research since the early days of the Internet. Tzanetakis and Cook (2002) addressed this problem with supervised machine learning approaches such as Gaussian Mixture model and k- nearest neighbour classifiers. They introduced 3 sets of features for this task categorized as timbral structure, rhythmic content and pitch content. Hidden Markov Models (HMMs), which have been extensively used for speech recognition tasks, have also been explored for music genre classification (Scaringella and Zoia, 2005; Soltau et al., 1998). Support vector machines (SVMs) with different distance metrics are studied and compared in Mandel and Ellis (2005) for classiying genre.

In Lidy and Rauber (2005), the authors discuss the contribution of psycho-acoustic features for recognizing music genre, especially the importance of STFT taken on the Bark Scale (Zwicker and Fastl, 1999). Mel-frequency cepstral coefficients (MFCCs), spectral contrast and spectral roll-off were some of the features used by (Tzane- takis and Cook, 2002). A combination of visual and acoustic features are used to train SVM and AdaBoost classifiers in Nanni et al. (2016).
With the recent success of deep neural networks, a number of studies apply these techniques to speech and other forms of audio data (Abdel- Hamid et al., 2014; Gemmeke et al., 2017). Representing audio in the time domain for input to neural networks is not very straight-forward because of the high sampling rate of audio signals. However, it has been addressed in Van Den Oord et al. (2016) for audio generation tasks. A common alternative representation is the spectrogram of a signal which captures both time and frequency information. Spectrograms can be considered as images and used to train convolutional neural networks (CNNs) (Wyse, 2017). A CNN was developed to predict the music genre using the raw MFCC matrix as input in Li et al. (2010). In Lidy and Schindler (2016), a constant Q-transform (CQT) spectrogram was provided as input to the CNN to achieve the same task.

This work aims to provide a comparative study between
1) The deep learning based models which only require the spectrogram as input and,
2) The traditional machine learning classifiers that need to be trained with hand-crafted features. We also investigate the relative importance of different features.

2. Data collection :

Audio Data: We have collected audio data using the youtube-dl download manager. The execution was carried out by coding a Python file with the latest changes in youtube-dl for audio samples. According to the research paper author, we were supposed to download 34 GB of .wav files for Data Processing. Initially, 40540 Files were expected to download, but due to computational and processing time, we have downloaded 25230 files with around 22.3 GB of data. Hence our dataset consists of 25230 audio .wav files.

The sample size varied differently, consisting of 7-Different types of music genres: Rock, Pop, Hip Hop, Techno, Rhythm Blues, Vocal and Reggae Music. Each of these Music Genres has a sufficient equivalent distribution of similar proportion without particular bias or dilution. Hence, we had around 25230 files ready for preprocessing for the ML model.

3. Historical Context

This involves understanding the development and evolution of Music Genre Classification methods over the course of time. Some key milestones and developments in this field are described below:

In the early days of music genre classification, manual methods were used. Experts and musicologists would categorize Music based on their subjective criteria such as instrumentation, vocal style, lyrical themes etc. These methods were limited by their subjectivity and lack of scalability.

- Feature-based classification: With the rise of computational methods, researchers had started exploring feature-based approaches for music genre classification. This approach allowed for automated genre classification based on objective audio features.
- Genre taxonomies and datasets: Prominent datasets, such as GTZAN and Million Song Dataset, provided labelled audio samples that were used for training and benchmarking genre classifiers.
- Machine learning algorithms: Support Vector Machines (SVM), k-Nearest Neighbors (k-NN), and Neural Networks have had a significant impact on music genre classification. These algorithms were applied to learn patterns and decision boundaries in the extracted audio features, improving the accuracy of genre classification.
- Deep learning approaches: Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) have been used in music genre classification. Deep learning models can automatically learn hierarchical representations from raw audio signals, leading to improved classification performance.
- Fusion of multi-modal data: In Our Project, we have used multiple models to get better and more accurate results. The fusion of multiple data modalities would enhance the music genre classification. Fusion provides a more comprehensive and holistic understanding of music genres.
- Evaluation Benchmarks: Usage of MIREX (Music Information Retrieval Evaluation eXchange) and the International Society for Music Information Retrieval (ISMIR) challenges, which provides standardized datasets and evaluation metrics.
- Real-world applications: Music genre classification techniques have found practical applications in various areas, including music recommendation systems, content-based music retrieval, playlist generation, and personalized Music streaming platforms. These applications leverage genre classification to enhance user experiences and provide tailored music recommendations.

Staying up-to-date with the latest literature and advancements in machine learning and music information retrieval will provide valuable insights into the historical context and ongoing progress in this area.

4. Case Studies :

1. "Music Genre Classification using Machine Learning Techniques" by Tzanetakis, G., & Cook, P. (2002):
2. "Automatic Music Genre Classification Using Deep Learning" by Hamel, J.-M., et al. (2010):
3. "Content-Based Music Genre Classification Using Deep Convolutional Neural Networks" by Choi, K., et al. (2016):
4. "Cross-genre Music Genre Classification using Convolutional Neural Networks" by Han, S., et al. (2017):
5. "Automatic Music Genre Classification with Transfer Learning" by Lee, K., et al. (2018):

These case studies showcase the usage of machine learning and deep learning techniques for music genre classification. Researchers have explored the use of Convolutional Neural Networks (CNNs), Support Vector Machines (SVM), and k-Nearest Neighbours (k-NN) algorithms. They have investigated various audio features and spectrogram representations of music to develop models that can automatically learn high-level representations. Transfer learning has been employed by leveraging pre-trained models from image datasets. Some studies have focused on cross-genre classification, addressing the challenge of classifying music into genres significantly different from the training genres. These were the case studies used for music-genre-classification.

6. Stakeholder analysis :

Genre Classification of a Song allows music enthusiasts to create a playlist of their favourite tracks. It also helps Music streaming services and platforms like Spotify, Apple Music, Soundcloud, etc., which provide recommendations to their users based on the genre of the songs they enjoy using various models. Automatic music genre classification is essential for organizations that sell Music or let users stream their content for a fee. The music database, as a result, would need to provide access to its users to search, retrieve, and Hence can make automatic recommendations to the listeners based on their taste.

The classification of the music data in the servers also increases the accessibility for its users. This ML Model serves the need to classify music samples into their respective genres. It helps to predict the genre using an audio file as its input. Automating the music classification seeks to make the selection of songs quick and less cumbersome. If one has to classify the songs or Music manually, one has to listen to many songs and then would be able to select the genre. This is time-consuming but also difficult. Hence, One such example is Shazam, and in this app, the music genre classification project has been used to find audio files with their genres and artists easily.

7. Regulation and Policies :

No specific regulations or policies apply directly to Music Genre Classification. However, there should be legal and ethical considerations related to data usage, and privacy of users, intellectual property of original creator/record labels may still be applicable.

Here are some general areas to consider for policy:

1. Data usage and privacy: Ensure the necessary rights and permissions to use the data with the explicit permission of regulatory bodies accordingly within regions. Respect any licensing agreements or copyright restrictions associated with the music files or datasets to be utilized.
2. Intellectual property: Ensure not violating any copyrights, trademarks, or other intellectual property protections when using music samples, excerpts, or other copyrighted materials during the training or evaluation of the model.
3. Data bias and fairness: Pay attention to potential biases in the training data that may affect the classification results. Ensure the training dataset is diverse and representative of music genres, cultures, and demographics to avoid bias in the classifier's outcomes.
4. User consent and transparency: Obtain the appropriate consent from users, which can also be in the form of Agreement Terms while signing up and should be transparent about the user's data will be used.

It's always a good practice to stay informed about the latest legal and ethical frameworks related to the usage of data and privacy-related rights. Additionally, it is important to consider the jurisdiction's specific requirements and regulations, as they may have guidelines or restrictions that could apply differently accordingly for users based on regions or other factors.

8. Existing projects and initiatives :

These are previous projects and initiatives related to our topic.Their Project Names with Research paper and author names are described as below:

Project: "Music Genre Classification with Transfer Learning"
Authors: Keunwoo Choi, George Fazekas, Mark Sandler
Research Paper: "Transfer Learning for Music Classification and Regression Tasks" (2017)

Project: "Genre Classification of Music using Deep Learning Techniques"
Authors: Kevin Cho, Euan Church, Geoffroy Peeters
Research Paper: "Automatic Musical Genre Classification of Audio Signals" (2014)

Project: "Deep Learning-based Music Genre Classification"
Authors: Jongpil Lee, Juhan Nam, Jangyeon Park, Kyogu Lee
Research Paper: "Automatic Tagging Using Deep Convolutional Neural Networks" (2017)

Project: "Music Genre Classification using Convolutional Neural Networks"
Authors: Oriol Nieto, George Tzanetakis, Juan Pablo Bello
Research Paper: "Music Genre Classification with Convolutional Neural Networks" (2016)

Project: "Music Genre Classification using Machine Learning Algorithms"
Authors: George Tzanetakis, Perry Cook
Research Paper: "Musical Genre Classification of Audio Signals" (2002)

These are just a few examples of projects and research papers in the field of Music Genre Classification ML, and there are many more active projects other than mentioned above related to this field of Music-genre-Classification.