Music Genre Classification Using Machine Learning

-Shubham Trivedi

DATE: 03-01-2022

1. Problem Statement

There are many recommendation systems for the music industry for the audience. but one thing they lack is that they do not provide recommendations of the song to a singer on the basis of his vocal capabilities. And I have tried to address that market opportunity

- ❖ What are Vocal Capabilities
  - Every person on this planet has the potential to be a singer as everyone is given the same set of vocal cords from birth for most of the cases. But to be a singer not only vocal cords but the skill to use them to their full capacity is also an art that is not practised by most people.
  - There are many Genres of Songs, Every genre is special and can be sung by your vocal cords but which is optimal?

2. Market/Customer/Business Need Assessment

- The music industry is famous for producing celebrities and millionaires worldwide, but one thing about the industry is that we seem to have a small number of music schools that cater to the growing number of people who choose music as their career. If you are interested in music, you should first think of ways to benefit from the thriving music industry, and one of the ways you can benefit from this career is to start your own music school.
- But when starting a new academy or being a music professor for years it is difficult to understand/predict the nature of a child and his vocal response to different types of songs you make him learn.
- But if we have some tool to predict the types of song suitable for a particular individual that becomes a whole lot easy and efficient to train on.
- I have made a Machine Learning model to do this prediction for you and your training center.

3. Target Specifications and Characterization

- Our main target for this will be students at music academies who are trying to learn music but don't know which type of songs will be suitable for them. Their teacher can assess their vocals on this model and then train students accordingly.
- Self-learners can take benefit from this product too.
- This model can also be seen as a fun way to see hidden capabilities for your natural untrained vocal cords.

4. External Search (information sources/references)

I have used the GTZAN Genre Collection dataset for this project Dataset can be found here:
http://opihi.cs.uvic.ca/sound/genres.tar.gz
The Dataset can be found on Marsyas.info. The dataset consists of 1000 audio tracks each 30 seconds long. It contains 10 genres, each represented by 100 tracks. The tracks are all 22050Hz Mono 16-bit audio files in .wav format.
Dataset Origin:
http://marsyas.info/downloads/datasets.html
MFCC Learning:
https://aircconline.com/sipij/V4N4/4413sipij08.pdf
Source:

- " Musical genre classification of audio signals " by G. Tzanetakis and P. Cook in IEEE Transactions on Audio and Speech Processing 2002.

5. Applicable Regulations

1. Above taken Dataset is obtained for the purpose of experiment , It cannot be used in a commercial setting.Thus, Acquiring the data to train the model is a must.for that Data protection and privacy regulations.
2. Government Regulations for small businesses.

6. Applicable Constraints

- Continuous data collection and maintenance
- Product's usage can be found not so reassuring as model's have low accuracy, Optimization is required in this case.
- Marketing the product to a Training center or a Software development company

7. Business Opportunity

1. Singing is a artform which is not a fortei of mst of the population.This product will help many aspiring artist or even general public who wants to expiercre the joy of singing.
2. Artforms are polished by the means of practice but to lay down a initial learning path is a very fast track process towards successful learning experience.
3. People when skillfully master a type of music that they are able to learn simply than the categories of songs that are harder for them is a main idea behind this product.
4. Teachers can know what kind of music is best suited for their new student based on the samples of their voice.

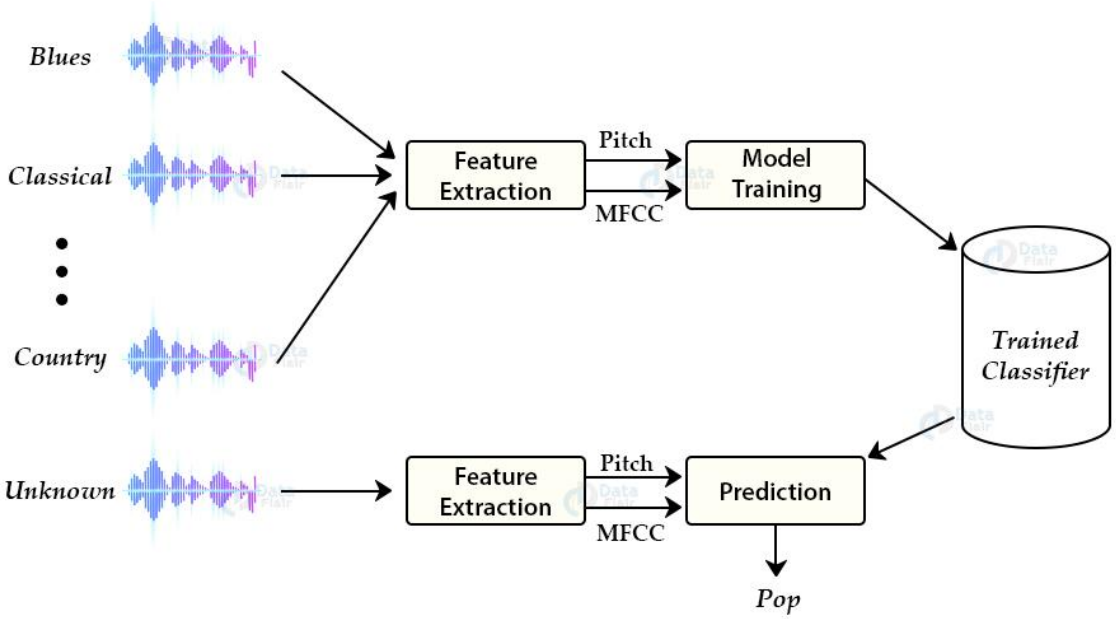
8. Concept Generation

1. There are lots of different kind of music which require even more different kind of features of vocals that are not same for any two people .Some features of voice are needed for specific type of music and other types require some different set of feature.For example, If someone has high bass and deep features in their natural voice then mastering the Blues or classical songs will be lot easier then learning pop songs as pop songs require somewhat high pitch voice.After learning the genre and mastering the songs by training their vocals users can easily jump into some not so suited genres as they have gained some experience controlling their vocal cords.So we are providing a suggestion of genre based on samples of audio files of the user's natural features.This product will require Machine learning algorithm in their toolbox which will provide whichkind of music will be best suite for the user to learn first.
   building a model would be a easy task but problem arise in optimizing it as pre-build models do not have very high accuracy rate and that will lead to incorrect suggestion of genres and ultimately user confidence in the product will be hamparred.

2. Model will take sample data to train on and extract the features of the songs from that model will train on

to better predict the results.After the training has been complete Classification of the user's vocal features will be carried out for which user given inputs which will be their voice samples will be fed to the classifier that we have trained.

3. features that have been extracted from user's input will be analysied and a genre of song will be given as output for the user to learn.Below,given figure explains the same .This product will use feature extraction and clustering algorithm at it's base and at the high level side of view recommendation system can be deployed for recommending materials and songs that are of the genre to the user to learn.



9. Concept Development

Input -> Feature Extration of audio samples -> Clustering Model -> Recommendation system -> Output

10. Final Report Prototype

The product takes the following functions to perfect and provide a good result

Back End

Model or Webapp Development: This must be done before releasing the service. A lot of manual supervised machine learning must be performed to optimize the automated tasks

- Performing EDA to realize the dependent and independent features.
- Algorithm training and optimization must be done to minimize overfitting of the model and hyperparameter tuning

Front End

- Different user interface: The user must be given many options to choose from in terms of parameters. This can only be optimized after a lot of testing and analysis of all the edge cases.
- Feedback system: A valuable feedback system must be developed to understand the customer's needs that have not been met. This will help us train the models constantly.

11. Code Implementation

This is a GitHub link:- https://github.com/Shubham-Trivedi/Shubham_trivedi_feynn_lab_Project1

12.Conclusion

There is not so many teaching systems made for beginners of singing and this product will be a step towards that and will help our user in understanding their own natural vocal capabilities.This is a integration of technical field through use of machine learning in hope to make learning to sing easier.