EMO PLAYER: Emotion Based Music Player

Hemanth P[1],Adarsh[1],Aswani C.B[1],Ajith P[1], Veena A Kumar[1]

[1] *Students, Department of Computer Science,Saintgits College Of Engineering, Pathamuttom, Kottayam, Kerala, India*
[1] *Assistant Proffesor, Department of Computer Science,Saintgits College Of Engineering, Pathamuttom, Kottayam, Kerala, India*

Abstract - *This project Emo player (an emotion based music player) is a novel approach that helps the user to automatically play songs based on the emotions of the user. It recognizes the facial emotions of the user and plays the songs according to their emotion. The emotions are recognized using a machine learning method Support Vector Machine (SVM)algorithm.*

SVM can be used for classification or regression problems. It finds an optimal boundary between the possible outputs. The training dataset which we used is Olivetti faces which contain 400 faces and its desired values or parameters. The webcam captures the image of the user.It then extract the facial features of the user from the captured image. The training process involves initializing some random values for say smiling and not smiling of our model, predict the output with those values, then compare it with the model's prediction and then adjust the values so that they match the predictions that were made previously.

Evaluation allows the testing of the model against data that has never been seen and used for training and is meant to be representative of how the model might perform when in the real world. According to the emotion ,the music will be played from the predefined directories

Key Words: SVM, Olivetti faces, Emotions, Songs, Machine learning, Training, Testing

1.INTRODUCTION

1.1 Background

Music plays a very important role in enhancing an individual‘s life as it is an important medium of entertainment for music lovers and listeners. In today‘s world, with ever increasing advancements in the field of multimedia and technology, various music players have been developed with features like fast forward, reverse, variable playback speed (seek & time compression),local playback, streaming playback with multicast streams and including volume modulation, genre classification etc. Although these features satisfy the user‘s basic requirements, yet the user has to face the task of manually browsing through the playlist of songs and select songs based on his current mood and behaviour.

1.2 Objectives

Project Emo player (an emotion based music player) is a novel approach that helps the user to automatically play songs based on the emotions of the user. It recognizes the facial emotions of the user and plays the songs according to their emotion. The emotions are recognized using a machine learning method Support Vector Machine(SVM)algorithm. The human face is an important organ of an individual‘s body and it especially plays an important role in extraction of an individual‘s behaviours and emotional state.The webcam captures the image of the user.It then extract the facial features of the user from the captured image. Facial expression categorized into 2, smiling and not smiling.

According to the emotion, the music will be played from the predefined directories.

1.3 PROBLEM DEFINITION

Using traditional music players, a user had to manually browse through his playlist and select songs that would soothe his mood and emotional experience. In today‘s world, with ever increasing advancements in the field of multimedia and technology, various music players have been developed with features like fast forward, reverse, variable playback speed (seek & time compression),local playback, streaming playback with multicast streams and including volume modulation, genre classification etc.

Although these features satisfy the user‘s basic requirements, yet the user has to face the task of manually browsing through the playlist of songs and select songs based on his current mood and behaviour. That is the requirements of an individual, a user sporadically suffered through the need and desire of browsing through his playlist, according to his mood and emotions.

1.4 PURPOSE

The main concept of this project is to automatically play songs based on the emotions of the user.

It aims to provide user-preferred music with emotion awareness. In existing system user want to manually select the songs, randomly played songs may not match to the mood of the user, user has to classify the songs into various

emotions and then for playing the songs user has to manually select a particular emotion. These difficulties can be avoided by using Emo Player (Emotion based music player). The emotions are recognized using a machine learning method Support Vector Machine (SVM)algorithm. SVM can be used for classification or regression problems. According to the emotion, the music will be played from the predefined directories.

1.5 SCOPE

Facial expressions are a great indicator of the state of a mind for a person. Indeed the most natural way to express emotions is through facial expressions. Humans tend to link the music they listen to, to the emotion they are feeling. The song playlists though are, at times too large to sort out automatically. It would be helpful if the music player was "smart enough" to sort out the music based on the current state of emotion the person is feeling. The project sets out to use various techniques for an emotion recognition system, analyzing the impacts of different techniques used.

By using Emo player we can easily play the songs according to the emotion of the user.

1.6 APPLICATION

- Automatically play song based on the emotion of the user.
- Act as a pluigin for website.
- Recommending for Youtube.
- Smart TV.
- Personal Assistant

1.7 LITERATURE REVIEW

Machine learning is a field of computer science that uses statistical techniques to give computer systems the ability to "learn" (i.e., progressively improve performance on a specific task) with data, without being explicitly programmed.

The name *machine learning* was coined in 1959 by Arthur Samuel. Evolved from the study of pattern recognition and computational learning theory in artificial intelligence, machine learning explores the study and construction of algorithms that can learn from and make predictions on data – such algorithms overcome following strictly static program instructions by making data-driven predictions or decisions, through building a model from sample inputs. Machine learning is employed in a range of computing tasks where designing and programming explicit algorithms with good performance is difficult or infeasible; example applications include email filtering, detection of network intruders or malicious insiders working towards a data breach, optical character recognition (OCR), learning to rank, and computer vision.

Machine learning is closely related to (and often overlaps with) computational statistics, which also focuses on prediction-making through the use of computers. It has strong ties to mathematical optimization, which delivers methods, theory and application domains to the field. Machine learning is sometimes conflated with data mining, where the latter subfield focuses more on exploratory data analysis and is known as unsupervised learning. Machine learning can also be unsupervised and be used to learn and establish baseline behavioral profiles for various entities and then used to find meaningful anomalies.

Within the field of data analytics, machine learning is a method used to devise complex models and algorithms that lend themselves to prediction; in commercial use, this is known as predictive analytics. These analytical models allow researchers, data scientists, engineers, and analysts to "produce reliable, repeatable decisions and results" and uncover "hidden insights" through learning from historical relationships and trends in the data. Effective machine learning is difficult because finding patterns is hard and often not enough training data are available; as a result, machine-learning programs often fail to deliver.

Our work is based and worked and it is run on python runtime library so we refer

[1] this reference gives python contribution for scientific area, and our player playes the music according to the type of emotion and it is based on support vector machine learning technique so we refer [2] which introduce some concepts of machine learning[3] We refer this to understand the language for data mining and machine learning.[4]This video link gives brief idea for emotion recognition in python.[5]This video link gives out brief idea of facial recognition in python. We need a dataset for our player to detect type of emotion , we need to generate it so we refer [6] for data set generation,[7] for emotion recognition with python and open CV.

Steps in Machine Learning

Machine learning is a field of computer science that gives computers the ability to learn without being programmed explicitly. The power of machine learning is that you can determine how to differentiate using models, rather than using human judgment. The basic steps that lead to machine learning and will teach you how it works are described below in a big picture:

1. Gathering data
2. Preparing that data
3. Choosing a model
4. Training
5. Evaluation
6. Hyper parameter tuning
7. Prediction.

Gathering Data:

Once you know exactly what you want and the equipment's are in hand, it takes you to the first real step of machine learning- Gathering Data. This step is very crucial as the quality and quantity of data gathered will directly determine how good the predictive model will turn out to be. The data collected is then tabulated and called as Training Data.

Data Preparation:

After the training data is gathered, you move on to the next step of machine learning: Data preparation, where the data is loaded into a suitable place and then prepared for use in machine learning training. Here, the data is first put all together and then the order is randomized as the order of data should not affect what is learned.

This is also a good enough time to do any visualizations of the data, as that will help you see if there are any relevant relationships between the different variables, how you can take their advantage and as well as show you if there are any data imbalances present. Also, the data now has to be split into two parts. The first part that is used in training our model, will be the majority of the dataset and the second will be used for the evaluation of the trained model's performance. The other forms of adjusting and manipulation like normalization, error correction, and more take place at this step.

Choosing a model:

The next step that follows in the workflow is choosing a model among the many that researchers and data scientists have created over the years. Make the choice of the right one that should get the job done.

Training:

After the before steps are completed, you then move onto what is often considered the bulk of machine learning called training where the data is used to incrementally improve the model's ability to predict.

The training process involves initializing some random values for say A and B of our model, predict the output with those values, then compare it with the model's prediction and then adjust the values so that they match the predictions that were made previously.

This process then repeats and each cycle of updating is called one training step.

Evaluation:

Once training is complete, you now check if it is good enough using this step. This is where that dataset you set aside earlier comes into play. Evaluation allows the testing of the model against data that has never been seen and used for training and is meant to be representative of how the model might perform when in the real world.

Parameter Tuning:

Once the evaluation is over, any further improvement in your training can be possible by tuning the parameters. There were a few parameters that were implicitly assumed when the training was done. Another parameter included is the learning rate that defines how far the line is shifted during each step, based on the information from the previous training step. These values all play a role in the accuracy of the training model, and how long the training will take.

For models that are more complex, initial conditions play a significant role in the determination of the outcome of training. Differences can be seen depending on whether a model starts off training with values initialized to zeroes versus some distribution of values, which then leads to the question of which distribution is to be used. Since there are many considerations at this phase of training, it's important that you define what makes a model good. These parameters are referred to as Hyper parameters. The adjustment or tuning of these parameters depends on the dataset, model, and the training process. Once you are done with these parameters and are satisfied you can move on to the last step.

Prediction:

Machine learning is basically using data to answer questions. So this is the final step where you get to answer few questions. This is the point where the value of machine learning is realized. Here you can Finally use your model to predict the outcome of what you want.

The above-mentioned steps take you from where you create a model to where you Predict its output and thus acts as a learning path.

2. System Analysis

2.1 Existing System

The features available in the existing Music players present in computer systems are as follows: i. Manual selection of Songs ii. Party Shuffle iii. Playlists iv. Music squares where user has to classify the songs manually according to particular emotions for only four basic emotions .Those are Passionate, Calm, Joyful and Excitement.

Using traditional music players, a user had to manually browse through his playlist and select songs that would soothe his mood and emotional experience .In today's world, with ever increasing advancements in the field of multimedia and technology, various music players have been developed

with features like fast forward, reverse, variable playback speed (seek & time compression),local playback, streaming playback with multicast streams and including volume modulation, genre classification etc.

Although these features satisfy the user's basic requirements, yet the user has to face the task of manually browsing through the playlist of songs and select songs based on his current mood and behaviour. That is the requirements of an individual, a user sporadically suffered through the need and desire of browsing through his playlist, according to his mood and emotions.

2.2 Limitation Existing System

- It requires the user to manually select the songs.
- Randomly played songs may not match to the mood of the user.
- User has to classify the songs into various emotions and then for playing the songs user has to manually select a particular emotion.

2.3 PROPOSED SYSTEM

Here wepropose a Emotion based music player(Emo Player).Emo player is an music player which play songs according to the emotion of the user.. It aims to provide user-preferred music with emotion awareness. Emo player is based on the idea of automating much of the interaction between the music player and its user. The emotions are recognized using a machine learning method Support Vector Machine(SVM)algorithm. In machine learning, support vector machines are supervised learning models with associated learning algorithms that analyse data used for classification and regression analysis. It finds an optimal boundary between the possible outputs. The training dataset which we used is Olivetti faces which contain 400 faces and its desired values or parameters. The webcam captures the image of the user. It then extract the facial features of the user from the captured image. The training process involves initializing some random values for say smiling and not smiling of our model, predict the output with those values, then compare it with the model's prediction and then adjust the values so that they match the predictions that were made previously. Evaluation allows the testing of the model against data that has never been seen and used for training and is meant to be representative of how the model might perform when in the real world. According to the emotion, the music will be played from the predefined directories.

2.4 Advantages of Proposed System

- Users don't want to select song manually.
- No need of playlist.
- Users don't want to classify the songs based on the emotions.

3 SOFTWARE AND HARDWARE REQUIREMENTS

Hardware Requirements:

The most common set of requirements defined by any operating system or software application is the physical computer resources, also known as hardware. The hardware requirements required for this project are:

- Intel i3
- 4GB RAM
- Webcam
- Speaker

Software Requirements:

Software Requirements deal with defining software resource requirements and pre-requisites that need to be installed on a computer to provide optimal functioning of an application. These requirements or pre-requisites are generally not included in the software installation package and need to be installed separately before the software is installed. The software requirements that are required for this project are:

- Python 2.7
- Open CV 3.1

4 GENERAL DESCRIPTION

4.1 Product Perspective

This project works on the basis of a dataset ..The dataset is an Olivetti faces containing of 400Faces with different emotion in it. This dataset works on the basis of the svm algorithm which classifies the data set into test and training dataset .Test dataset helps to give out scores of the learning dataset or its accuracy whereas Training dataset helps to particularly helps to seek linear and non linear data after which it is classified and when camera is invoked certain facial features are captured and it is then used to find the training dataset being classified, this helps to select a particular type of emotion (happy/sad).then according to that emotion music will be played.

4.2 Product Functionality

- Training Dataset

In machine learning, the study and construction of algorithms that can learn from and make predictions on data is a common task. Such algorithms work by making data-driven predictions or decisions, through building a mathematical model from input data.

The data used to build the final model usually comes from multiple datasets. In particular, three data sets are

commonly used in different stages of the creation of the model.

The model is initially fit on a **training dataset**, that is a set of examples used to fit the parameters (e.g. weights of connections between neurons in artificial neural networks of the model. The model (e.g. a neural net or a naive Bayes classifier) is trained on the training dataset using a supervised learning method (e.g. gradient descent or stochastic gradient descent). In practice, the training dataset often consist of pairs of an input vector and the corresponding *answer* vector or scalar, which is commonly denoted as the *target*. The current model is run with the training dataset and produces a result, which is then compared with the *target*, for each input vector in the training dataset. Based on the result of the comparison and the specific learning algorithm being used, the parameters of the model are adjusted. The model fitting can include both variable selection and parameter estimation.

Successively, the fitted model is used to predict the responses for the observations in a second dataset called the validation dataset. The validation dataset provides an unbiased evaluation of a model fit on the training dataset while tuning the model's hyperparameters (e.g. the number of hidden units in a neural network). Validation datasets can be used for regularization by early stopping: stop training when the error on the validation dataset increases, as this is a sign of overfitting to the training dataset.

- SVM Algorithm
-

In machine learning, support vector machines (SVMs, also support vector networks) are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis. Given a set of training examples, each marked as belonging to one or the other of two categories, an SVM training algorithm builds a model that assigns new examples to one category or the other, making it a non-probabilistic binary linear classifier (although methods such as Platt scaling exist to use SVM in a probabilistic classification setting). An SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall. More formally, a support vector machine constructs a hyperplane or set of hyperplanes in a high- or infinite-dimensional space, which can be used for classification, regression, or other tasks like outliers detection. Intuitively, a good separation is achieved by the hyperplane that has the largest distance to the nearest training-data point of any class (so-called functional margin), since in general the larger the margin the lower the generalization error of the classifier.

- Webcam module

A webcam is a video camera that feeds or streams its image in real time to or through a computer to a computer network. When "captured" by the computer, the video stream may be saved, viewed or sent on to other networks via systems such as the internet, and emailed as an attachment. When sent to a remote location, the video stream may be saved, viewed or on sent there. Unlike an IP camera (which connects using Ethernet or Wi-Fi), a webcam is generally connected by a USB cable, or similar cable, or built into computer hardware, such as laptops.

The term "webcam" (a clipped compound) may also be used in its original sense of a video camera connected to the Web continuously for an indefinite time, rather than for a particular session, generally supplying a view for anyone who visits its web page over the Internet. Some of them, for example, those used as online traffic cameras, are expensive, rugged professional video cameras. This camera usually invoked after the dataset is initialized. This the captures the essential facial features which is then searched in hyper plane to find out the particular emotion(happy/sad) and plays the music according to the emotion. Cross Validation is a very useful technique for assessing the performance of machine learning models. It helps in knowing how the machine learning model would generalize to an independent data set.

4.3 User Characteristics

Since it is a security system version 1.0, the user need not know what the code actually does. During requirement gathering, the user need to specify some important information such as details about a particular email where he/she want to receive the information about the laptop location once it is stolen. It is considered that the user do not have the basic knowledge of operation behind the system and to have access to it.

4.4 Functional Requirements

Functional requirements are statement of services the system should provide, how the system should react to particular inputs and how the system should behave in particular situation.:

- The dataset train by support vector classifier
- Machine learns support vector classification using support vector machine.
- Learn and identify image capture by web cam.

4.5 Non Functional Requirements

Non functional requirements define system properties and constraints it arises through user needs, because of budget constraints or organizational policies, or due to the external

factors such as safety regulations, privacy registration and so on.

Non functional requirements are:

- Reliability
- Maintainability
- Portability
- Extensibility
- Reusability
- Simplicity.
- Resource Utilization

5. CONCLUSIONS

This project has been developed to give us great advancement in the field of machine learning technology. EMO player fulfills to sort out the music based on the emotions of the user such as whether it is happy or sad . So, totally our work aims to develop a player which is based on user need and it helps to revive in case of free time or leisure time if we want to hear music based on our current situation.

REFERENCES

[1] J. Bergstra, O. Breuleux, F. Bastien, P. Lamblin, R. Pascanu, G. Desjardins, J. Turian, D. Warde-Farley, Y. Bengio, "Theano: a CPU and GPU math expression compiler", Proceedings of the Python for Scientific Computing Conference (SciPy), Jun. 2010.

[2] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, "Scikit -learn: Machine learning in Python", Journal of Machine Learning Research, vol. 12, pp. 2825-2830, 2011.

[3] L. Buitinck, G. Louppe, M. Blondel, F. Pedregosa, A. Mueller, O. Grisel, V. Niculae, P. Prettenhofer, A. Gramfort, J. Grobler, R. Layton, J. VanderPlas, A. Joly, B. Holt, G. Varoquaux, "API design for machine learning software: experiences from the scikit-learn project", ECML PKDD Workshop: Languages for Data Mining and Machine Learning, pp. 108-122, 2013.

[4]https://www.youtube.com/watch?v=eIIfb5D3HO8

[5]https://www.youtube.com/watch?v=4W5M-YaJtIA

[6]https://thecodacus.com/opencv-face-recognition-python-part1/

[7]http://www.paulvangent.com/2016/04/01/emotion-recognition-with-python-opencv-and-a-face-dataset/

[8]https://pythonprogramming.net/haar-cascade-face-eye-detection-python-opencv-tutorial