Intro to Machine Learning

Complex Engineering Problem

Submitted by

2018-EE-018
2018-EE-020
2018-EE-108
2018-EE-151

19,December 2021

University of Engineering and Technology
Lahore, Pakistan

Contents

List of Figures

1 Abstract

Speech Emotion Recognition is the act of attempting to recognize human emotion and affective states from speech. This is capitalizing on the fact that voice often reflects underlying emotion through tone and pitch. This is also the phenomenon that animals like dogs and horses employ to be able to understand human emotion. SER is tough because emotions are subjective and annotating audio is challenging.

2 Introduction

The voice signal is the quickest and the most regular technique for correspondence between people. This reality has motivated scientists to consider voice a quick and productive strategy for cooperation among human and machine. Nonetheless, this necessitates that the machine ought to have the adequate insight to perceive human voices. Since the late fifties, there has been colossal examination on voice acknowledgment, which suggests to the most common way of changing over the human voice into a grouping of words. Be that as it may, regardless of the incredible headway gave in voice acknowledgment, we are still a long way from having a characteristic collaboration among man and machine in light of the fact that the machine doesn't comprehend the enthusiastic condition of the speaker.

This has presented a moderately late exploration field, in particular voice feeling acknowledgment, which is characterized as removing the enthusiastic condition of a speaker from their voice. It is accepted that vocal feeling acknowledgment can be utilized to separate valuable semantics from voice, and henceforth, works on the exhibition of vocal acknowledgment frameworks.

Vocal feeling acknowledgment is especially helpful for applications which require normal man–machine collaboration, for example, web motion pictures and PC instructional exercise applications where the reaction of those frameworks to the client relies upon the distinguished feeling.[1]

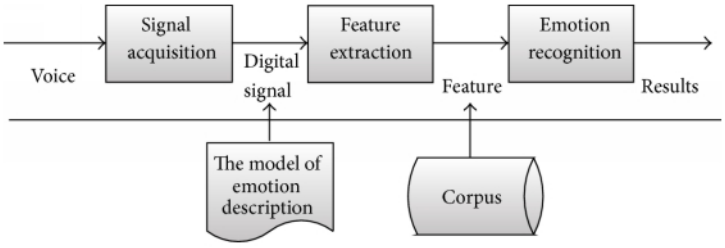
It is additionally helpful for in-vehicle board framework where data of the psychological condition of the driver might be given to the framework to start his/her security. It very well may be likewise utilized as a suggestive instrument for advisors.

It could be likewise valuable in programmed interpretation frameworks in which the enthusiastic condition of the speaker assumes a significant part in correspondence between parties. In airplane cabins, it has been observed that speech emotion acknowledgment frameworks prepared to pushed speech accomplish preferred execution over those prepared by typical speech.[2]

Speech feeling acknowledgment has additionally been utilized in call place applications and versatile correspondence. The fundamental target of utilizing discourse feeling acknowledgment is to adjust the framework reaction after distinguishing disappointment or irritation in the speaker's voice.

The assignment of speech feeling acknowledgment is extremely trying for the accompanying reasons. To start with, it isn't clear which speech highlights are generally amazing in recognizing feelings. The acoustic inconstancy presented by the presence of various sentences, speakers, talking styles, and talking rates adds another problem in light of the fact that these properties straightforwardly influence the majority of the normal extricated speech

highlights like pitch, and energy shapes. [3]Besides, there might be more than one saw feeling in a similar expression; every feeling compares to an alternate piece of the verbally expressed expression. Moreover, it is truly challenging to decide the limits between these parts. Another difficult issue is that how a specific feeling is communicated by and large relies upon the speaker, their way of life and climate.



3 Methodology

3.1 Data set collection

A total of 1344 sentences dataset is collected. As per 12 speakers recorded 4 emotions i.e. happy, angry, sad and neutral with 7 dialogues each. So we go as 12*4*7*4 = 1344

3.2 Information Retrieval

Librosa Python package is used for audio analysis. Librosa is basically used when we work with audio data like in music generation and Automatic Speech Recognition. It provides the building blocks necessary to create the music information retrieval systems. The data and sampling rate can be obtained from the library. For audio categorization it resemples our 48000Hz mono sounds to 22050Hz.

3.3 Feature Extraction

Features are extracted from the sound file.Because Audio feature extraction is a necessary step in audio signal processing as it removes unwanted noise and balances the time-frequency ranges by converting digital and analog signals.

3.4 Dataset Division

Dataset is divided such that 75percent is training data and 25percent is test dataset.

3.5 Dataset Training

Dataset is trained on these algorithms i.e. Logistic Regression (LR), Random Forest (RF), Support vector Machine (SVM)

The details of these algorithms are as given below:

Logistic Regression

Regression analysis to conduct when the dependent variable is binary 0 or 1 is Logistic Regression.
Catagorical dataset of dependent variables, we cannot use linear regression we will use logistic regression. Classification is done in it.Using Sigmoid function this function convert independent variables into expression of probability that ranges between 0 and 1 with respect to dependent variable.Whatever the

predictions they will be converted into probabilities that have range between 0 and 1. Binary classification is the strength of Logistic Regression. Logistic regression plots real values in the form of yes or no (1 or 0) . Some points are plotted over the cutoff line and rest below the cutoff line. And the data points that are on the cutoff line they are not classifiable.
There are a few things that should be kept in mind while using logistic regression listed below:

- Whatever data set we are using it should not be having any missing value this data set should be complete.
- 30-50 Data points for each output
- 60-100 Binary logistic regression data points.

Applications

- Fraud Detection
- Disease Diagnosis
- Emergency Detection
- Email spam or non spam

Accuracy Results of LR

```
[64] acc_lr = []
     clf_lr = make_pipeline(StandardScaler(), LogisticRegression(solver='saga', max_iter=200, random_state=0))
     clf_lr.fit(x_train, y_train)
     acc_lr.append(clf_lr.score(x_test, y_test))
     print("The average accuracy for Logistic Regression is {}%".format(np.average(acc_lr)*100))

     The average accuracy for Logistic Regression is 82.14285714285714%
```

Random Forest Algorithm

Random forest is a supervised learning algorithm.It builds a forest of an ensemble trees that are trained by Bagging method.The combination of learning models increase overall result is general idea of Bagging method.
random forest builds multiple decision trees and merges them together to get a more accurate and stable prediction. We can use it for two type of problems:
1. Regression
2. Classification

The hyperparameters for RF algo are same as that of decision tree(greedy search method) or bagging method. This algo adds randomness in the model.We can use random threshold as well this is another advantage of this algorithem.Anyone who needs to build the model in short duration of time can use this algo as it requires less time. It is very hard to challenge this algo's performance.[4] RF is simple fast and flexible tool but it also has certain limitations that are listed below:
1.Large number of trees reduce the speed of algorithem very slow.
2. It is ineffective for real time predictions.
Accuracy Results of RF Algorithem

```
[59] clf=RandomForestClassifier(n_estimators = 200, criterion = 'entropy',max_depth=200,random_state=0)
     clf.fit(x_train,y_train)
     y_pred1=clf.predict(x_test)
     print("Accuracy for RF",accuracy_score(y_test,y_pred1)*100)

     Accuracy for RF 88.98809523809523
```

Support Vector Machine(SVM) algorithm

A supervised machine learning algorithm that can be used for both classification or regression challenges is known as Support Vector Machine(SVM) algorithm.
SVM lies in supervised learning used for classification and regression analysis. It is of two types:

- Quadrilateral
- Circle

A data set that is having quadrilateral as well as circle thus training dataset is labeled as SVM follows supervised learning. This data set is given to trainer that builds model. When the model is ready it us passed through testing phase where prediction is done whether it will go in quadrilateral on in circle.

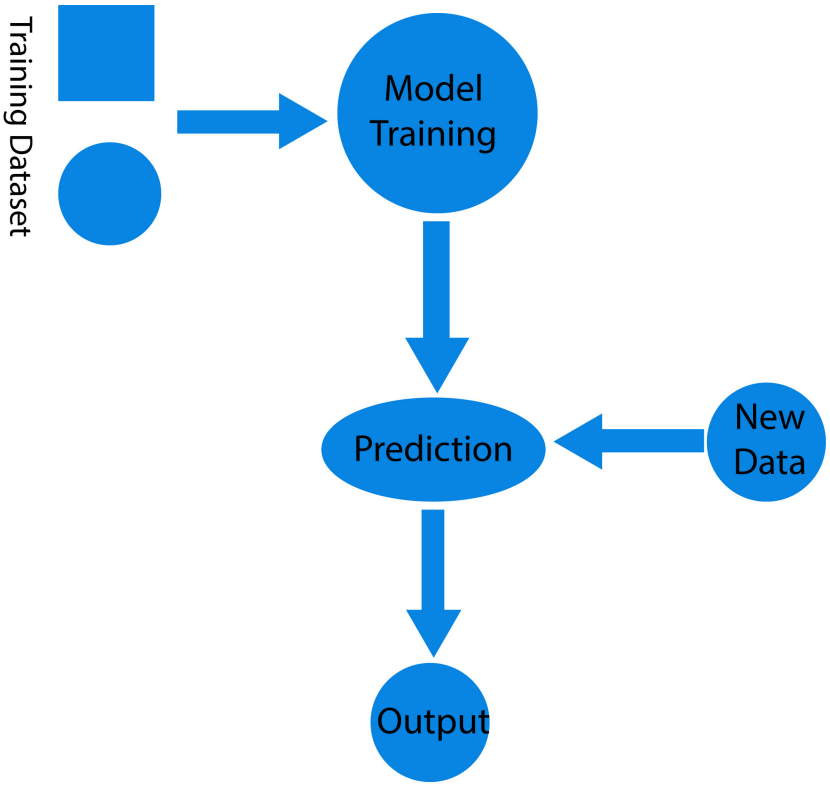
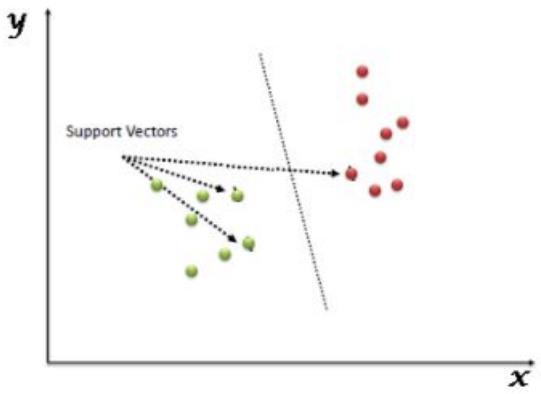


Figure 1: SVM Block Diagram

If our model is accurate and we are giving a quadrilateral data set and the output we obtained through should be quadrilateral (i.e it belongs to Quadrilateral Class). Results in xy-plane:



We want to separate the dataset so that we will draw a line between them that line is called decision boundary (hyperplane). This decision boundary decides that this new data belongs to circle class or quadrilateral class. But the question is how to get this line? We will see the resemblance of data set of circle with quadrilateral and the one that is similar we shall draw a line that joins it with the other class. Similarly a line will be drawn through quadrilateral's data set that will be parallel to the first obtained line. By joining these two lines together we get two distances. Determine the margin. Margin plays an important role in deciding the hyperplane. Margin will decide which hyperplane will exist. The dataset can be linear or non linear. Which type of SVM will depend on the type of dataset.[5] Accuracy Results of SVM Algorithm

```
[34] acc_svm = []
     clf_svm = SVC(kernel='linear')
     clf_svm.fit(x_train, y_train)
     acc_svm.append(clf_svm.score(x_test, y_test))
     print("The average accuracy for SVM classifier is {}%".format(np.average(acc_svm)*100))

     The average accuracy for SVM classifier is 82.73809523809523%
```

4 Experimental Setting

Sampling Frequency = 48000Hz mono sound

Random Forest Parameters

Estimators = 200
Criterion = Entropy
max-depth=200
random-state=0

Logistic Regression Parameters

solver='saga'
max-iter=200
random-state=0

State Vector Machine Parameters

kernel='linear'

5 Conclusion

We used three algorithms as tabled below to train our model. Their accuracies are as

Sr.no	Algorithm	Accuracy
1	RF	88.9880
2	LR	82.1428
3	SVM	82.7380

In our dataset Random forest's accuracy is best beacuse it is great with high dimensional data since we are working with subsets of data. It is faster to train than decision trees because we are working only on a subset of features in this model, so we can easily work with hundreds of features. These accuracies can be improved by increasing hidden layers and getting more data set.

6 References

[1]https://www.analyticsvidhya.com/blog/2017/09/understaing-support-vector-machine-example-code/
[2]https://towardsdatascience.com/understanding-random-forest-58381e0602d2
[3]https://www.investopedia.com/terms/t/t-test.asp
[4]https://machinelearningmastery.com/logistic-regression-for-machine-learning/