ISSN 2348-1196 (print)
International Journal of Computer Science and Information Technology Research ISSN 2348-120X (online)
Vol. 3, Issue 2, pp: (714-716), Month: April - June 2015, Available at: www.researchpublish.com

A Novel Approach for Automatic Mood Classification of Indian Popular Music

Shalini Malik

Department of Computer Science, Maharishi Dayanand University, Rohtak, India

Abstract: **We most often choose to listen to a song or music which best fits our mood at that instant. In spite of this strong correlation, most of the music software's presents today are still devoid of providing the facility of mood-aware play-list generation. This increase the time music listeners take in manually choosing a list of songs suiting a particular mood or occasion, which can be avoided by annotating songs with the relevant emotion category they convey. We contribute by making an effort to build a system for automatic identification of mood underlying the audio songs by mining their spectral, temporal audio features. Our focus is specifically on Indian Popular Hindi songs. We have analyzed various data classification algorithms in order to learn, train and test the model representing the moods of these audio songs and developed an open source framework for the same.**

Keywords: **Random forest, Bagging, Music, Domain knowledge, Mood identification.**

1. MINING MOOD FROM AUDIO FEATURES

Decision trees have been used for classification due to:

a) Does not require any domain knowledge or parameter setting

b) Decision trees can handle high dimensional data.

c) Their representation of acquired knowledge in tree form is intuitive

d) The learning and classification steps of decision tree induction are simple and fast.

1.1 Random Forest Classification:

In order to improve the classification accuracy, ensemble methods like bagging and boosting have been proved quite productive. Ensemble methods use a combination of models of a series of k learned classification models , M_1, M_2,..., M_k, with the aim of creating an improvised model in terms of classification accuracy. Our work makes use of the "Bagging" approach also called as "Boot-strap aggregation". In this method, bootstrap samples of data-sets are created by randomly sampling the features and data instances from the given training set with replacement. These samples are then independently and simultaneously used for training and learning classifier models separately for each sample. Finally, the classification is done by considering the maximum of the votes taken from all the models learnt

Benefits of the approach:-

- Random forests readily handle large number of classifiers.
- They are faster to train and evaluate as compared to other comparable approaches
- Random forests exhibit stronger resistance to over-training and thus over fitting.
- Separate Cross-validation is unnecessary in case of Random forests since it is already taken care at the time of forest building.

- Random forests generally have similar accuracy as Support Vector Machines, Neural Networks although Random forests have shown much better performance in case of huge and high-dimensional data-sets

1.2 Bagging of random forests:

In this work we present an additional hierarchy of ensemble by generating an ensemble of Random Forests using bootstrap aggregation also know as Bagging. For growing Random Trees, the randomly sampled data attributes are split on the basis of "Gini Index" which has shown better results when working with CART trees. Gini Index, basically measures the impurity of the data set. The approach used in Algorithm has not only shown a rise in accuracy of classification of music data-sets as compared to traditional Random forest approach, but also has shown a consistent better performance as compared to other classification techniques

2. MOOD IDENTIFICATION SYSTEM

The Mood Identification system is the main engine which would help identify the mood of given music or audio files. This system is designed as an open source software system. The system would generally be a part of the back-end in most of the applications whose result can be used by the application layer on top to utilize the information in the required way. The system has two-fold objectives as mentioned below:-

1. The system should have a provision of analyzing music files and learn the classifier models associated with them

2. It should be able to predict the class of mood that a particular audio file or music belongs to.

2.1 Working of the system:

1. Input music files

2. Audio Preprocessing

3. Audio feature extracting

4. Mood classifier learner (using mood model)

5. Mood detector (using mood model)

6. Mood is the output

3. EVALUATION METRICS

1. Receiver Operating Characteristic- : It shows the trade-of between the true positive rate and the false positive rate. It is a two-dimensional plot with vertical axis representing the true positive rate and horizontal axis representing the false positive rate. A model with perfect accuracy will have an area of "1".

The area under the ROC curve is a measure of the accuracy of the model. It ranks the test tuples in decreasing order: the one that is most likely to belong to the positive class appears at the top of the list .The closer to the diagonal line (i.e., the closer the area is to 0.5), the less accurate is the model.

2. Confusion Matrix - The columns of the confusion matrix represent the predictions, and the rows represent the actual class. Correct predictions always lie on the diagonal of the matrix.

3. Recall is a metric that gives a percentage of how many of the actual class members the classifier correctly identified. (FN + TP) represent a total of all minority members.

$$\text{Recall} = TP/ (TP + FN)$$

4. Precision- It gives us the total the percentage of how many of a particular class instances as determined by the model or classifier actually belong to that particular class. (TP + FP) represents the total of positive predictions by the classifier

$$\text{Precision} = TP/(TP+ FP)$$

5. F-Measure*:* It is a harmonic mean of Precision and Recall. We can say that it is essentially an average between the two percentages. It really simplifies the comparison between the classifiers.

F- Measure = 2/(1/rec + 1/prec).

4. CONCLUSION

We successfully experimented with the task of mapping audio features of Indian Popular Music with respective moods with the top precision ranging in between 75% and 81% with respect to Fmeasure and 70% to 75% precision measure. The best accuracy w.r.t. area under ROC was observed in the range 0.91 to 0.94 which seems quite satisfactory. The Bagging of Random Forest approach thus performed much better as compared to not just other decision tree based algorithms but other classification algorithms as well.

REFERENCES

[1] Tsunoo, E., Akase, T., Ono, N., Sagayama S., (2010), "Musical mood classification by rhythm and bass-line unit pattern analysis", Proceedings of IEEE In- ternational Conference on Acoustics, Speech and Signal Processing (ICASSP).

[2] Attribute-Relation File Format, http://www.cs.waikato.ac.nz/ml/weka/arf.html

[3] Doris Baum, (2006), "Emomusic - Classifying music according to emotional",Proceedings of the 7th Workshop on Data Analysis (WDA), Kosice, Slovakia.

[4] Breiman, L., (1996), "Bagging predictors". Machine Learning 24(2), 123140.

[5] Capurso, A., Fisichelli, V. R., Gilman, L., Gutheil, E. A., Wright, J. T., (1952),"Music and Your Emotions", Liveright Publishing Corporation.

[6] Chia-Chu Liu, Yi-Hsuan Yang, Ping-Hao Wu, Homer H. Chen, (2006), "Detecting and classifying emotions in popular music", JCIS Proceedings'.

[7] Dalibor Mitrovic, Matthias Zeppelzauer, Horst Eidenberger, (2007), "Analysis of the Data Quality of Audio Descriptions of Environmental Sounds", Journal of Digital Information Management, 5(2):48.

[8] Trung-Thanh Dang, Kiyoaki Shirai, (2009), "Machine Learning Approaches for Mood Classification of Songs toward Music Search Engine", International Conference on Knowledge and Systems Engineering.