Supervised Machine Learning for Urban Sound Classification

A fshankaleem Sr.[1], Assistant Professor[2], *Dr.I.Santi Prabha Professor*[3],
[1]*Dept of ECE, MJCET, Hyderabad, TS, INDIA*
[2]*Dept. Of ECE, JNTUCE, JNTUK, Kakinada, A.P, INDIA*

Abstract- Machine learning is a significant developing area for almost all researches today. Attaining knowledge from empirical data is the backbone of machine learning. The knowledge isderived by changing either structure or parameters of an algorithm model or both in order to bring an improvement in its expected performance on future data. These changes have been accommodated to accomplish one of artificial intelligence taskswhich can be decision making, recognition, learning, diagnosis, prediction, control or planning, etc. Different approaches are used to learn from data such as Naïve Bayes, j48Decision Tree, Random Forest.

In this paper we will discuss classifyingthe databy designingvarious models and doing their comparativeanalyses Urban sound dataset and let us have a brief look at its recognition accuracy pattern with respect to various models.

I. INTRODUCTION

Machine learning is a significant developing area for almost all researches today. Attaining knowledge from empirical data is the backbone of machine learning. The knowledge is derived by changing either structure or parameters of an algorithm model or both in order to bring an improvement in its expected performance on future data. These changes have been accommodated to accomplish one of artificial intelligence tasks which can be decision making, recognition, learning, diagnosis, prediction, control or planning, etc.Different approaches are used to learn fromdata such as NaïveBayes, j48Decision Tree, Random Forest.

In this paper we will discuss learning from experimental data by using various models and their comparison study. I have used URBAN SOUND 8k dataset and let us have a brief look at its recognition accuracy pattern with respect to various models.

A. CLASSIFICATION OVERVIEW

Classification of audio signals include severalrelevant tasks for instance source identification, automatic speech recognition, automatic musictran scription, sentiment/ emotion recognition, music/speech/environmental sound segmentation. Common machinelearning techniques are applicable in the above respective sub fields.

It is the process by which we automatically assign an individual item to one of these number of categories or

C.MACHINE LEARNING ALGORITHMS

The following are a brief overview of some Machine learning algorithms:

NAÏVE BAYES: Naive Bayes is among one of the most simple and powerful algorithms for classification based on Bayes' Theorem with an assumption of independence among predictors. Naive Bayes model is easy to build and particularly useful for very large data sets. There are two parts to this algorithm:

- **Naive**
- **Bayes**

classes based on its characteristics. The complexity lies infinding an appropriate relationship between features and classes. In our case:

1. The items are audiosignals (example: sounds, tracks, excerpts)
2. Their characteristics are the features we extract from them (MFCC).
3. The classes (sound environment) fit the problem definition.

B.Block diagram of Audio Recognition System

An input audio signal is the initial requirement given to the system, where the raw input of data is being processed and, in this segment, the voiced and unvoiced sounds are being segregated. Also, the unwanted background noises are eliminated, and this segment is known as Preprocessing via many methods such as Filtering, Short term energy, etc. Followed by whichis the Feature Extraction technique whose main criteria is to extract certain parameters of interest to the users. Generally, these parameters are distinctive factors that help distinguish one audio wave from another for instance loudness, pitch, timbre, etc.

It's a small amount of data from the voice that can be used to represent each speaking source. Further the revised audio signals undergo a training and testing phase. Here various Machine Learning models can be used to match the unknown set of audio data to the trained set of data via algorithms. This is known as feature matching.

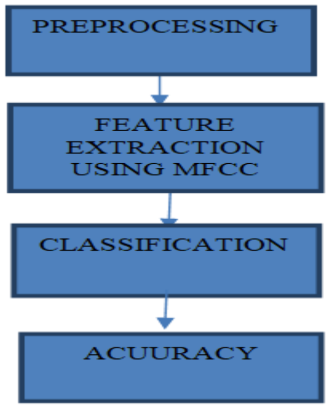


Fig.1: WORK FLOW

The Naive Bayes classifier assumes that the presence of a feature in a class is unrelated to any other feature. Even if these features depend on each other or upon the existence of the other features, all of these properties independently contribute to the probability that a particular fruit is an apple or an orange or a banana and that is why it is known as "Naive".

In statistics and probability theory, Bayes' theorem describes the probability of an event, based on prior knowledge of conditions that might be related to the event. It serves as a way to figure out conditional probability.

Given a Hypothesis (H) and evidence (E), Bayes' Theorem states that the relationship between the probability of the hypothesis before getting the evidence, P(H), and the probability of the hypothesis after getting the evidence, P(H|E), is:

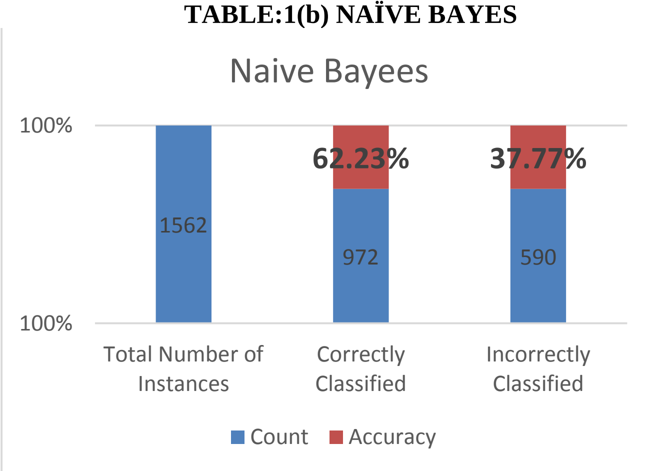
$$P(H|E) = \frac{P(E|H) \,.\, P(H)}{P(E)} \qquad \ldots\ldots\ldots\ldots\ldots 1$$

For this reason, P(H) is called the prior probability, while P(H|E) is called the posterior probability. The factor that relates the two, P(H|E)/P(E), is called the likelihood ratio. Using these terms, Bayes' theorem can be rephrased as: "The posterior probability equals the prior probability times the likelihood ratio."

TABLE:1(a) NAÏVE BAYES

CLASIFFIERS	Summary	Mfccc(40)	
j48	Total number of instances	1562	
	Correctly classified	1159	74.19%
	Incorrectly Classified	403	25.80%
	Statistics		
	Time to build model	0.38sec	
	Mean Absolute Error	0.054	
	Root Mean Square Error	0.021	

TABLE:1(b) NAÏVE BAYES



J48 DECISION TREE:

Decision Tree Algorithm is to find out the way the attributes-vector behaves for a number of instances. Also, on the bases of the training instances the classes for the newly generated instances are being found .This algorithm generates the rules for the prediction of the targetvariable. With the help of tree classification algorithm, the critical distribution of the data is easily understandable. J48is an extension of ID3. The additional features of J48 are accounting for missing values, decision trees pruning, continuous attribute value ranges, derivation of rules, etc.

The WEKA tool provides a number of options associated with tree pruning. In case of potential over fitting pruning can be used as a tool for précising. In other algorithms the classification is performed recursively till every single leaf is pure, that is the classification of the data should be as perfect as possible. This algorithm it generates the rules fromwhich particular identity of that data is generated. The objective is progressivelygeneralization of a decision tree until it gains equilibrium of flexibility and accuracy.

Steps in the Algorithm:

(i) In case the instances belong to the same class the tree represents a leaf so the leaf is returned by labelling with the same class.

(ii) In case the instances belong to the same class the tree represents a leaf so the leaf is returned by labelling with the same class.

(iii) The potential information is calculated for every attribute, given by a test on the attribute. Then the gain in information is calculated that would result from a test on the attribute.

(iv) Then the best attribute is found on the basis of the present selection criterion and that attribute selected for branching

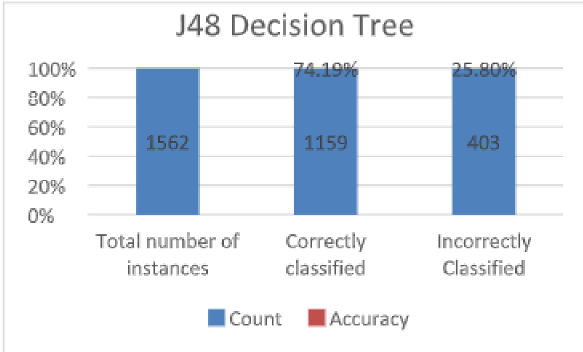


TABLE:2 (a) J48 DECISION TREE

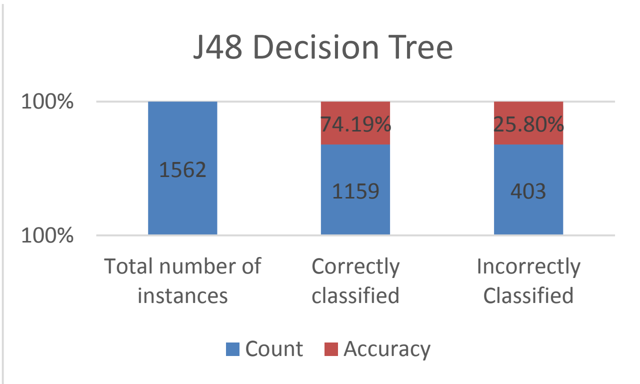


TABLE:2 (b) J48 DECISION TREE

Random Forest:

Random Forest is a flexible, easy to use machine learning algorithm that produces, even without hyper-parameter tuning, a great result most of the time. It is also one of the most used algorithms, because it's simplicity and the fact that it can be used for both classification and regression tasks. Random Forest is a supervised learning algorithm. it creates a forest and makes it somehow random.The forest it builds, is an ensemble of Decision Trees, most of the time trained with the "bagging" method. The general idea of the bagging method is that a combination of learning models increases the overall result.To say it in simple words: Random forest builds multiple decision trees and merges them together to get a more accurate and stable predictionRandom Forest has nearly the same hyperparameters as a decision tree or a bagging classifier.

Random Forest adds additional randomness to the model, while growing the trees. Instead of searching for the most important feature while splitting a node, it searches for the best feature among a random subset of features. This results in a wide diversity that generally results in a better model.Another great quality of the random forest algorithm is that it is very easy to measure the relative importance of each feature on the prediction.Sklearn provides a great tool for this, that measures a features importance by looking at how much the tree nodes, which use that feature, reduce impurity across all trees in the forest. It computes this score automatically for each feature after training and scales the results, so that the sum of all importance is equal to 1.

Through looking at the feature importance, you can decide which features you may want to drop, because they don't contribute enough or nothing to the prediction process. This is important, because a general rule in machine learning is that the more features you have, the more likely your model will suffer from overfitting and vice versa.

Advantages

- An advantage of random forest is that it can be used for both regression and classification tasks.
- Random Forest is also considered as a very handy and easy to use algorithm, because it's default hyperparameters often produce a good prediction result.

Disadvantages

- The main limitation of Random Forest is that a large number of trees can make the algorithm to slow and ineffective for real-time predictions. In general, these algorithms are fast to train, but quite slow to create predictions once they are trained.
- A more accurate prediction requires more trees, which results in a slower model.
- In most real-world applications the random forest algorithm is fast enough, but there can certainly be situations where run-time performance is important and other approaches would be preferred.

CLASIFFIERS	Summary	Mfccc(40)	
RANDOM FOREST	Total number of instances	1562	
	Correctly classified	1462	93.59%
	Incorrectly Classified	100	6.04%
	Statistics		
	Time to build model	0.86sec	
	Mean Absolute Error	0.068	
	Root Mean Square Error	0.147	

TABLE:3(a) RANDOM FOREST

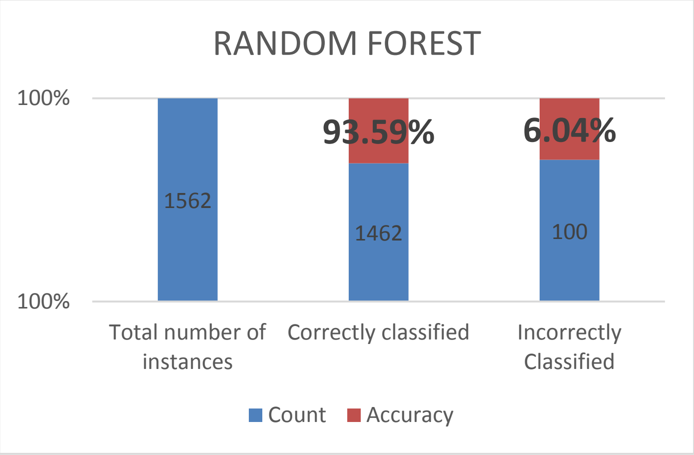
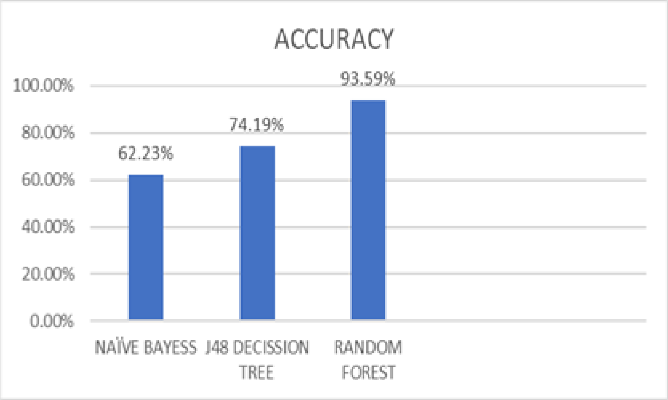


TABLE:3(b) RANDOM FOREST

*A. **Results***: The accuracy results obtained for the various classifiers are as follows, and observations declare that Random Forest out performs the remaining classifier able to increase the Accuracy Rate. Experimental Results have Shown Significant improvement of accuracy up to 93.59%.

TABLE:4 ACCURACY



B.Challenges:

- Separating urban sounds from background noise.
- Discontinuity in speech.
- Homonyms can be a challenge in ASR systems.
- Overlapping of sounds/speech.

II. CONCLUSION

Through this Research work we intent to extract features and classify them from an Urban Sound Data Set that is Urbansound 8k to enable city agencies Field, Military systems, Community Services, Educational Purposes, Command and Control eg: Systems which can automatically recognize digits/data spoken into a telephone. Random Forest has been.

III. ACKNOWLEDGMENT

I sincerely thank and acknowledge all the authors and researchers for their contributed work and landmark in the field of Machine Learning and specifically Speech Recognition and the related technology that gives us the window to think and reflect.

IV. REFERENCES

[1]. L.Breiman, “Pasting small votes for classification in large databases and on-line”, Machine Learning, 36(1), 85-103, 1999

[2]. SoniaSuuny , David Peter S , K. Poulose Jacob, “Performance of Different Classifiers In Speech Recognition”, 2013 IJRET

[3]. Justin Salmon, Christopher Jacoby, Juan Pablo Bello”Music and Audio research Laboratory, New York University,Centre of Urban Science and Progress”,2014,USA

[4]. Huan Zhou, Ying Song , Haiyan Shu,”Smart Energy and Environment Cluster Institute for Infocomm Research”, IEEE Region Conference(TENCON), Malaysia,November 5-8,2017

[5]. N. R. Fatahillah, P.Suryati and C. Haryawan, "Implementation of Naive Bayes classifier algorithm on social media (Twitter) to the teaching of Indonesian hate speech," 2017 International Conference on Sustainable Information

[6]. N.Albarakati and V. Kecman, "Fast neural network algorithm for solving classification tasks: Batch error back-propagation algorithm," 2013 Proceedings of IEEE Southeastcon, Jacksonville, FL, 2013, pp. 1-8

[7]. A. B. Kandali, A. Routray and T. K. Basu, "Emotion recognition from Assamese speeches using MFCC features and GMM classifier," TENCON 2008 - 2008 IEEE Region 10 Conference, Hyderabad, 2008, pp. 1-5