International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9, Issue-7, May 2020

Classification of Musical Instruments using SVM and KNN

S. Prabavathy, V. Rathikarani, P. Dhanalakshmi.

Abstract: Automatic classification of musical instruments is a challenging task. Music data classification has become a very popular research in the digital world. Classification of the musical instruments required a huge manual process. This system classifies the musical instruments from a several acoustic features that includes MFCC, Sonogram and MFCC combined with Sonogram. SVM and kNN are two modeling techniques used to classify the features. In this paper, to simply musical instruments classifications based on its features which are extracted from various instruments using recent algorithms. The proposed work compares the performance of kNN with SVM. Identifying the musical instruments and computing its accuracy is performed with the help of SVM and kNN classifier, using the combination of MFCC and Sonogram with SVM a high accuracy rate of 98% achieve in classifying musical instruments. The system tested sixteen musical instruments to find out the accuracy level using SVM and kNN.

Keywords: Musical Instruments Classification, Mel-frequency cepstral coefficients (MFCC), Feature extraction, k-Nearest Neighborhood (kNN), Sonogram, Support Vector Machine (SVM).

I. INTRODUCTION

Music is a means of communication and is an integral part of many human activities. Most ancient cultures used music for celebrations, worship, healing and preserving their stories. The fundamental characters of the tones produced are pitch, loudness, duration and timbre. One of the basic functions of a musical instrument is to produce tones of the desired pitch [1]. Classification [2] is a process of a class label assigning to the given observation. Using signal processing techniques, the raw audio signal are computed from the features in the musical analysis systems. MFCC is used to construct a feature vector. It is adequate and to determine the best way to perform their transformation. This shows that MFCCs not only suitable for speech and also for music modeling [3]. SVM is used to consider a simple way in binary classification. SVM is a conceptual learning machine that will learn from a set of training and try to generalize and build correct predictions on fresh data [4]. The performance of kNN is much better when all data having the same scale. It stores the training dataset which it uses as its representation.

Revised Manuscript Received on May 30, 2020.
* Correspondence Author
S. Prabavathy*, Research Scholar, Department of Computer and Information Science, Department of Computer Science and Engineering, Annamalai University, Chidambaram, Tamilnadu, India.
V. Rathikarani, Assistant Professor, Department of Computer Science and Engineering, Annamalai University, Chidambaram, Tamilnadu, India.
P. Dhanalakshmi, Professor, Department of Computer Science and Engineering, Annamalai University, Chidambaram, Tamilnadu, India.

This algorithm makes the predictions by calculating the similarities between the input samples and the each training instance [5]. In [6] SVM and kNN are used for the classification of musical instruments; MFCC, Sonogram and MFCC combined with Sonogram are used for feature extraction. Musical instruments are classified according to this method which is used to make sound. The musical instruments were classified by their material. A traditional classification was called as eight sounds and produced by eight different kinds of materials. They are stone, metal, thread, gourd, skin, earth, bamboo and wood.
There are many types of musical instruments, some of the musical instruments are used for classification such as banjo, cello, guitar, mandolin and violin from a string instrument, bass clarinet, bassoon, clarinet, contrabassoon, flute and saxophone from woodwind instrument, frenchhorn, trombone, trumpet and tuba from brass instrument and piano from keyboard instrument. 1284 music samples were collected from various musical instrument databases. The block diagram of this proposed work is shown is fig. 1

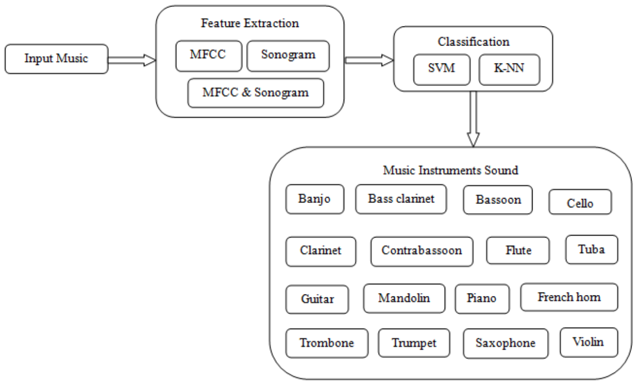

Fig. 1The block diagram of Musical Instruments Classification.

II. LITERATURE REVIEW

For each isolated word, features were extracted and trained successfully by SVM. Audio content is characterized using the features of Sonogram in [7]. In [8] musical instruments like western were classified with HOS features were combined with MFCC. The kNN classifier is used for hierarchical classification. The proposed methods in [9] like kNN, GMM, ANN and dropout ANN used to classify European orchestral musical instruments. The proposed algorithm [10] various types of music genres classified using the support vector machine (SVM) that can perform binary classification. In [11] the audio and video data classified into any of the five following classes such as advertisement, news, cartoon, movie and songs using Mel frequency cepstral coefficients (MFCC) with colored histogram features were extracted from images in video clips that used for visual features.

In [12] speech emotion recognized using k-Nearest Neighbor (kNN) and Gaussian Mixture Model (GMM) models to recognize six emotional categories such as happy, neutral, angry, surprised, sad and fearful. Emotion like 'happy' gives more accuracy whereas 'fear' and 'surprise' emotion gives the lowest rate. The computation speed of the kNN classifier is fast when compared to GMM classifier. Using the Minimum Distance Classifier in [13] the musical instruments were classified. By calculating the FRR, the results analyzed. The proposed algorithm [14] uses Auto Associative Neural Network (AANN) based on features using Sonogram to recognized speech.

III. FEATURE EXTRACTION

Mel-frequency cepstral coefficients (MFCC)

MFCC is used to construct a feature vector. Great amount of researchers already inspecting for what purposes of MFCCs are adequate and it determines the finest way to perform those transformations. This shows that MFCCs not only suitable for speech and also for music modeling. MFCCs are widely accepted in the field of speech recognition, human speech and voice identification [15]. In [16] musical genre is classified based on the short speech signals by using MFCC as a feature extraction method. MFCC is used to characterize audio content by calculating the features. In [17] AANN is used to segment and classify the audio signal automatically by using LPC, LPCC and MFCC as a feature extractor. The extraction of MFCC from the music signal is shown in Fig 2.

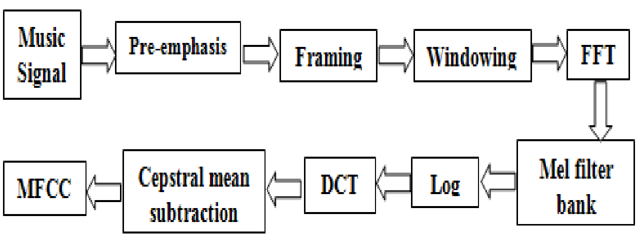

Fig.2 Extraction of MFCC from music signal

Sonogram

Sonogram is a current incarnation of feature set, music data is sampled at 22 kHz, mono format. In [7] the features of each isolated word extracted and the model is trained effectively. To characterize the audio content, Sonogram is calculated as features. SVM classifier has been used to recognize speech by learn from the training data.

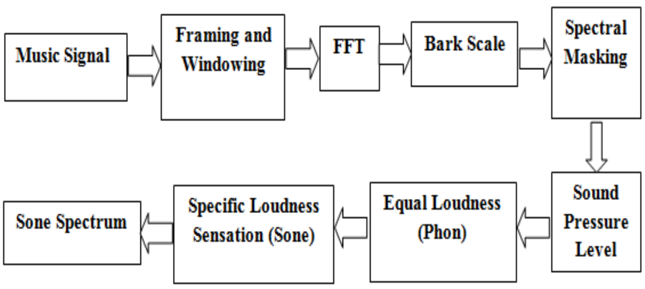

Fig. 3 Sonogram Feature Extraction

In [14] spoken word converted into text using AANN model using Sonogram as acoustic feature. Feature vectors from Sonogram were given as input. VAD is used to separate distinct words from the continuous speeches. The block diagram of Sonogram feature extraction is shown in fig. 3.

IV. CLASSIFIERS

Support vector machine (SVM)

Support vector machines are used for nonlinear regression and pattern classification. It constructs the linear model, with non-linear boundaries based on the support vectors to approximate the result function. The data are separated linearly, the linear machines are trained for the optimal hyperplane which split the data without miscalculation and into the most distance between the hyperplane and the nearby training points. These training points are closest to the optimal separating hyperplane are called the support vectors. SVM architecture shows in Fig. 4.1.

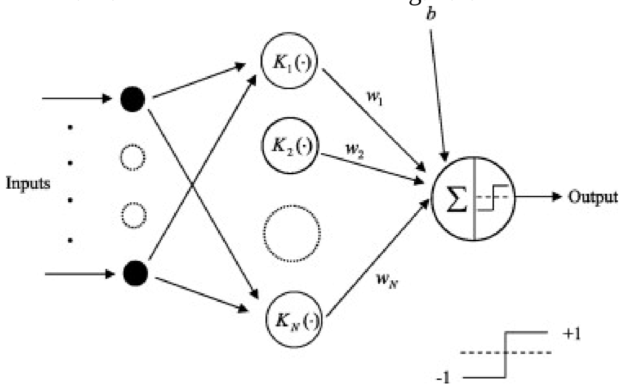

Fig. 4.1 Architecture of SVM

K-Nearest Neighbor (kNN)

kNN algorithm is very easy and very efficient. The output data is calculated while the class with the peak frequency as of the k-most related instances. For each instance the votes for their class and the class with a large amount of votes is taken as prediction [5]. Fig. 4.2 shows the process of K-Nearest Neighbor (kNN).

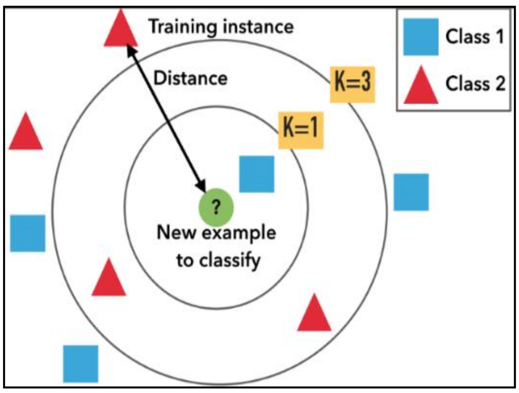

Fig. 4.2 K-Nearest Neighbor

V. PROPOSED WORK

In this proposed work, the music signals are preprocessed to remove noise before feature extraction, 39 MFCC features, 22 Sonogram features and 61 MFCC and Sonogram features were extracted from each and every music samples. In python, a package named LibROSA which is used to convert sound to FFT, using the audio data the MFCC & Sonogram plot feature extraction done. In training, the saved features are loaded as input to the classification algorithm such as SVM and kNN. Using the confusion matrix the Recall, Precision, F-Score and Accuracy were calculated.

The trained classifiers are saved as a model and inference to that model which is used to classify the class. In SVM, linear kernel is used to classify the class that one class compares the other classes like one versus rest. In kNN, the k value initialize, to classify the class iterate from class 1 to the nearest five neighbors and calculated by the highest vote. In testing, the trained features were given to the classifier algorithm as input such as kNN and SVM to classify the musical instruments. The block diagram for music instrument classification shows in fig. 5.

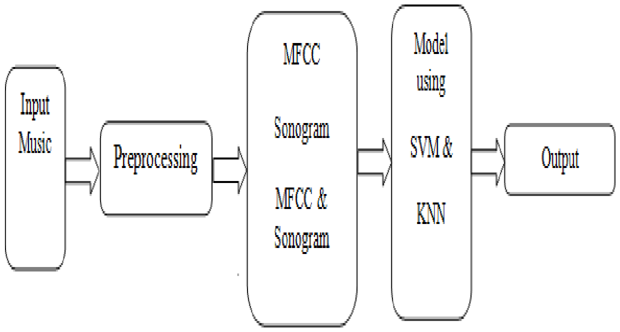

Fig. 5 Block Diagram for Musical Instrument Classification

VI. EXPERIMENTAL RESULTS

Dataset

The musical instruments sound data taken from the RWC database, musicbrainz.org, MINIM-UK musical instrument database, IRMAS: a dataset for instrument recognition in musical audio signals and NSynth dataset. The duration of music samples were range from 1 sec. to 2 min. and 1284 music samples were collected from sixteen different musical instruments from four different families such as string, woodwind, keyboard and brass. 70% of musical samples were trained and rest of data used for testing.

6.1 Classification using MFCC with SVM and kNN

In the proposed work, the features were extracted by MFCC and the classifiers SVM and kNN used to classify the musical instruments. 98% of accuracy yield by this proposed system in classifying the musical.

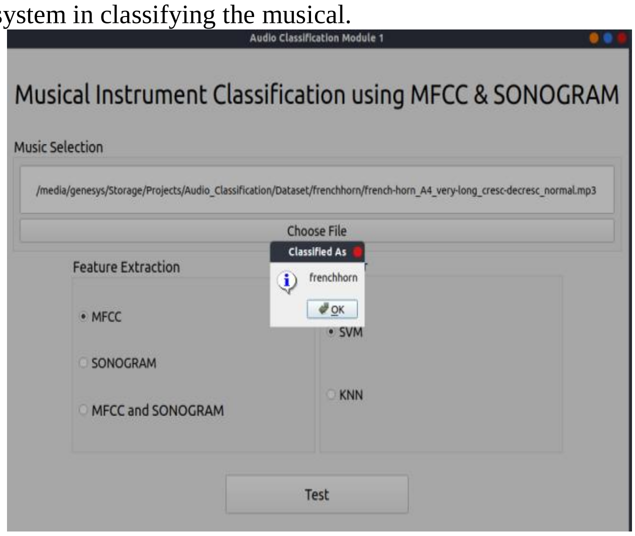

Fig. 6.1 (a) Classification using MFCC with SVM

Fig. (6.1(a)) shows the classification of musical instruments using MFCC with SVM and fig. (6.1 (b)) shows the musical instrument classification using MFCC with kNN.

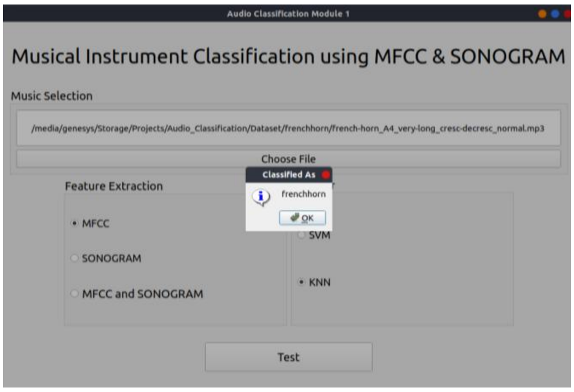

Fig. 6.1 (b) Classification using MFCC with kNN

6.2 Classification using Sonogram with SVM and kNN

In the proposed work the features were extracted by Sonogram and the classifiers SVM and kNN used to classify the musical instruments.

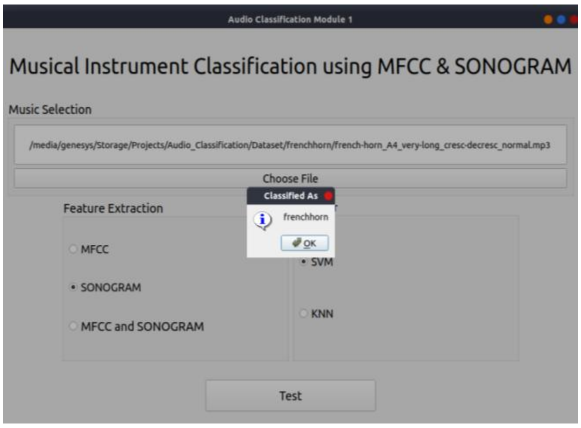

Fig. 6.2. (a) Classification using Sonogram with SVM

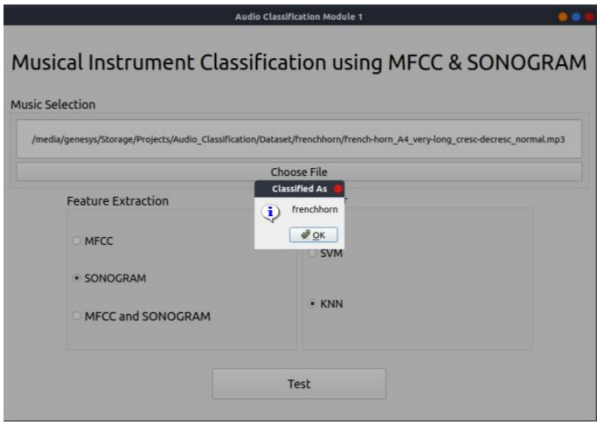

Fig. 6.2 (b) Classification using Sonogram with kNN

The accuracy of the proposed system kNN gives 95% and 97% for SVM in classifying the musical instrument. Fig. (6.2 (a)) shows the musical instrument classification using Sonogram with SVM and fig. (6.2 (b)) shows the musical instrument classification using Sonogram with kNN.

6.3 Classification using MFCC and Sonogram with SVM and kNN

In the proposed work the features were extracted by MFCC combined with Sonogram and the classifiers SVM and kNN used to classify the musical instruments.

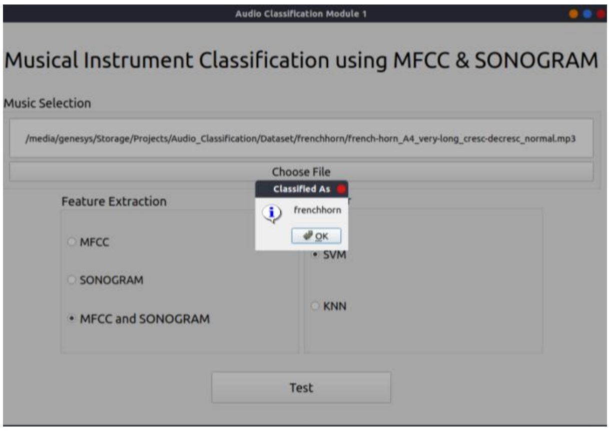

Fig. 6.3 (a) Classification using MFCC & Sonogram with SVM

The accuracy of the proposed system kNN gives 98% and 99% for SVM in classifying the musical instrument. Fig. (6.3(a)) shows the musical instrument classification using MFCC and Sonogram with SVM and fig. (6.3 (b)) shows the musical instrument classification using MFCC and Sonogram with kNN.

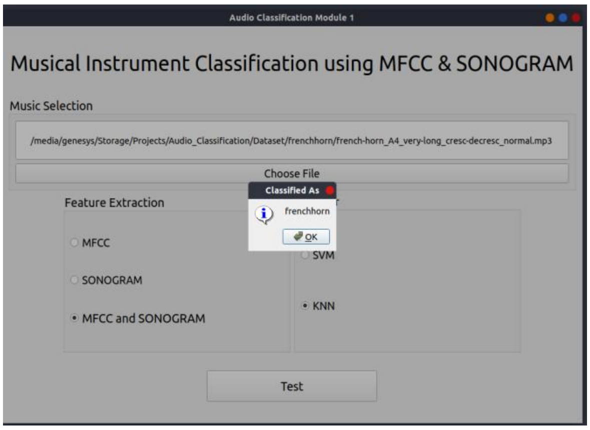

Fig. 6.2 (b) Classification using MFCC & Sonogram with kNN

VII. PERFORMANCE MEASURES

Different numbers of sample set from 16 musical instruments are taken for the proposed system. The overall performance of F-Score, precision, recall and accuracy of musical instruments shows in table 1. The accuracy of every instrument is calculated using the confusion matrix. The musical data be trained and tested effectively; precision, recall, F-Score and accuracy of musical instruments are

Precision = TP / (TP + FP)
Recall = TP / (TP + FN)
Accuracy = (TP + TN) / (TP + TN + FP + FN)

Table 1 Precision, Accuracy, Recall and F-Score of Musical Instruments

Features	Classifiers	Precision (%)	Recall (%)	F-Score (%)	Accuracy (%)
MFCC	SVM	89.59	83.87	86.63	97.53
	kNN	88.38	86.19	87.27	98.22
Sonogram	SVM	88.44	81.76	84.96	97.98
	kNN	74.52	70.75	72.58	95.68
MFCC & sonogram	SVM	96.21	94.12	95.15	99.29
	kNN	87.11	84.78	85.92	98.17

The accuracy of SVM is good when compare to kNN in this proposed work. Fig. 7 shows the overall accuracy comparison of the proposed work.

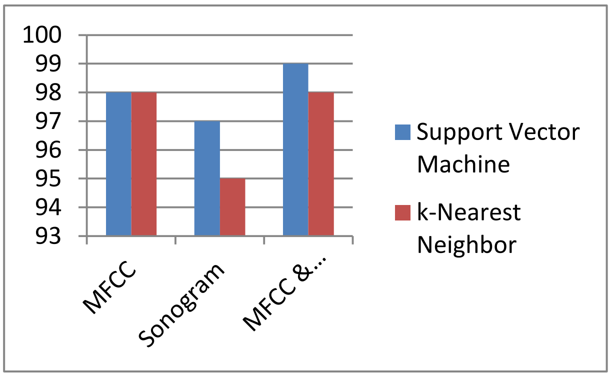

Fig. 7 Accuracy comparison of proposed work

VIII. CONCLUSION

In this paper, automatic classification of musical instrument system has been proposed using SVM and kNN where MFCC, Sonogram and their combination are calculated as features to classify the musical instruments. Experimental results show that, the musical data collected from various sources for classification. The combination of MFCC and Sonogram with SVM gives high accuracy of 98% and Sonogram with kNN gives the least accuracy of 95%. In this proposed work the performance of the combination of MFCC and Sonogram with SVM is better when compared to Sonogram with kNN. In future the musical instruments will classify using deep learning.

REFERENCES

1. Musical Sound, Instruments, and Equipment, By Panos Photinos
2. Signal Processing Methods for Music Transcription edited by Anssi Klapuri, Manuel Davy
3. Computing with Instinct: Rediscovering Artificial Intelligence by Yang Cai
4. Learning with Support Vector Machines, By Colin Campbell, Yiming Ying
5. Master Machine Learning Algorithms: Discover How They Work and Implement Them From Scratch, By Jason Brownlee
6. The Science of String Instruments, edited by Thomas D. Rossing.
7. R. Thiruvengatanadhan, Speech Recognition using Sonogram and SVM , 2018 IJRAR January 2019, Volume 06, Issue 1 www.ijrar.org (E-ISSN 2348-1269, P- ISSN 2349-5138)
8. [8] Kazi F I, Bhalke D G, Musical Instrument Classification using Higher Order Spectra and MFCC, International Conference on Pervasive Computing (ICPC), -1-4799-6272-3/15/$31.00, 2015 IEEE
9. *Miroslav MALÍK, Richard ORJEŠEK*,The Comparison of Selected Audio Features and Classification Techniques in the Task of the Musical Instrument Recognition, M. Malík, R. Orješek, Musical Instrument Recognition, Poster 2016, Prague May 24.

10. Lei-Ting Chen, Ming-Jen Wang, Chia-Jiu Wang, and Heng-Ming Tai, Audio Signal Classification Using Support Vector Machines, J. Wang et al. (Eds.): ISNN 2006, LNCS 3972, pp. 188 – 193, 2006. Springer-Verlag Berlin Heidelberg 2006.
11. K. Subashini, S. Palanivel, V. Ramalingam Audio-Video based Classification using SVM and AANN, *International Journal of Computer Applications (0975 – 8887) Volume 44– No.6, April 2012*
12. Rahul B. Lanjewar, Swarup Mathurkar, Nilesh Patel, Implementation and Comparison of Speech Emotion Recognition System using Gaussian Mixture Model (GMM) and K – Nearest Neighbor (kNN) Techniques *Rahul B. Lanjewar et al. / Procedia Computer Science 49 (2015) 50 – 57*
13. Sayantani Nandi, Madhura Banerjee, Parangama Sinha, Jayati Ghosh Dastidar, SVM Based Classification of Sounds from Musical Instruments using MFCC Features, Sayantani Nandi *et al*, International Journal of Advanced Research in Computer Science, 8 (5), May-June 2017,2144-2147
14. R. Thiruvengatanadhan, Speech Recognition Using Sonogram And AANN, International Journal of Innovative Research in Advanced Engineering (IJIRAE) ISSN: 2349-2163 Issue 01, Volume 6 (January 2019).
15. Computing with Instinct: Rediscovering Artificial Intelligence, edited by Yang Cai.
16. Gursimran Kour, Neha Mehan, Music Genre Classification using MFCC, SVM and BPNN, *International Journal of Computer Applications (0975 – 8887) Volume 112 – No. 6, February 2015.*
17. P. Dhanalakshmi, S. Palanivel and M. Arul, Automatic Segmentation Of Broadcast Audio Signals Using Auto Associative Neural Networks, ICTACT Journal On Communication Technology, December 2010, VOL.1, ISSUE: 04, ISSN: 2229– 6948 (ONLINE).

AUTHORS PROFILE

S. Prabavathy, M.C.A,. She is currently pursuing Ph.D in Annamalai University, Chidambaram, Tamilnadu, India. Her research interests are machine learning and deep learning.

V. Rathikarani, B.E., M.E., Ph.D (CSE). Pattern Classification, Computer Networks, Medical Image Processing are her interests. She is Assistant Professor in Annamalai University, Chidambaram, Tamil Nadu, India.

P. Dhanalakshmi, B.E., M.Tech., M.B.A., Ph.D. Audio and Speech Signal Processing, Data Mining, Pattern Recognition and Software Engineering are her interests. She is Professor in Annamalai University, Chidambaram, Tamil Nadu, India.