International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

A SYSTEM FOR BIRD CLASSIFIER USING MACHINE LEARNING APPROACH

[1] Sirisha Godavari , [2] M. Srivani, [3] M. Harshavardhan Reddy, [4] M. Shiva Prasad Reddy, [5] S. Vinay kumar

[1]Assistant Professor in Department of CSE Sri Indu College Of Engineering And Technology
Sirisha.godavari@gmail.com
[2,3,4,5] UG Scholars Department of CSE Sri Indu College Of Engineering And Technology

Abstract:

Bird Classifier is a system which will use machine learning approach and allows users to record bird sounds and identify them. The System will help bird species to be identified by taking input only the sound of that bird. The framework also offers tools for translating such annotations into datasets that can be used to train a computer to identify a species ' presence or absence. Dataset recorded are preprocessed to remove unwanted noise and divide useful sounds in frames so that it can act as an input to classifier. The system was tested through different algorithms and the algorithm that gave best results was chosen for implementation. System uses audio features like MFCC,Mel- Spectra etc. This system will use different algorithms such as KNN, Random Forest, Multi layer perceptron, Bayes in classification of birds species.

I INTRODUCTION

Recently, technology has developed a lot, especially in the field of Machine Learning (ML), which is useful for reducing human work. In the field of artificial intelligence, ML integrates statistics and computer science to build algorithms that get more efficient when they are subject to relevant data rather than being given specific instructions [1][2][3]. Machine learning is commonly used in diverse fields to solve difficult problems that cannot be readily solved based on computer approaches [4][5][6]. Recently, these advances in machine learning have helped a lot with sound classification, and sound recognition has shown to be a strong value in automating these tasks [7]. To say it another way, birds can make two basic sounds: Call and song [8][9]. While this approach is timeconsuming, machine learning approaches may also be useful in establishing differentiating between the different species, even after that since it is done on a species of birds that are still not thought to be discernable [10]. However, machine learning's usage of bird classification has only been examined for a small number of species or mannequin processing on the assumption that it can be applied in the real world only through numerical simulation or hand recordings [11]. The results have proven unpractical for ecologists but can be

useful for many people of a wide variety of professions [12]. When the classification rate study is extended to more organisms that are currently existing, the findings may vary greatly [13][14]. The features used to identify and classify birds can be organized in two ways: First, bird species can be compared, and second, all birds can be identified based on a handful of specific features. In the sound classification of birds, the static noise is removed because it makes it difficult to hear the bird calls until the signal is filtered and boost the volume. The study of the bird's specific sound categories such as joyful, sad, gentle, grating, and quiet to discover a lot of additional information about it. Machine learning algorithms, examine them to decide which strategies are the most effective at identifying birds [15-20]. The more often used audio feature— (MFCC). Melfrequency cepstral coefficients (MFCCs), MFC is composed of individual values that add up to a unit vector. The form of the vocal tract expresses itself in the time-band continuum, and the function of (MFCCs) is to faithfully capture **it**

II LITERATURE SURVEY

V. Morfi et al. [21] presented NIPS4Bplus which is the first annotated, typographically enriched bird song dataset. The NIPS4Bplus dataset and tags used for the 2013 bird song classification challenge, as well as newly acquired temporal annotations, make up NIPS4Bplus. They have comparative data on the recordings, as well as their species-specific tags and temporal annotations. A. Pareta et al. [22] The MC-LS-VM classifier was calculated with an RBF kernel function with seven input parameters given a class accuracy features a rating of 85.43% According to them, their claims have had the best results to date and are therefore more successful to date. K. Ko et al. [23] utilized the pre-trained neural network offers innovative solutions for fine categorization of animal species based on their sound signals using pre-trained CNNs, and a new self-attention model well-suited for acoustics. I. Lezhenin et al. [24] proposed the LSTM model outperforms a range of current implementations and is more accurate and reliable than the previous model CNN. L. Nanni et al. [25] analyzed CNN model using two collections of animal data: one on the feline audio files and the other on bird recordings. They also devised a way to locate the centers of the spectra through their analysis of their preexposed dots. When experimenting with their proposal, results show that it outperforms other approaches on other types of data as well Since the use of LSTM networks is highly effective in learning temporal dependencies W. Xu et al. [26] implemented a CNN design, in which three convolutions had to be done in parallel using three different filter lengths Observation showed that their method was successful, obtaining a high degree of accuracy in real-world environments. Erhan Akbal [27] achieved a 90.25% prediction accuracy rate with his proposed system. In his Mehyadin et al.;

AJRCOS, 9(4): 1-11, 2021; Article no.AJRCOS.68530 3 study, he suggests a cognitive, lightweight, highly reliable, and low bandwidth form of online expansion. By the end of the experiment, it was shown that this approach works. M. M. M. Sukri et al. [28] used the bird's sound classification system employs artificial neural networks (ANN). This work has been completed and can provide useful information on the different bird types. Using the automated environmental sound classification (or, ESC) it is possible to foresee the kind of sounds that will be made. C. Chalmers et al. [29] obtained bird songs are sampled with a Mel-band filter bank cephalometer to collect their vocal parameters analyzed with a multilayer perceptron to determine whether they belong to one species or another. Their proposed approach yielded positive results with a sensitivity of 0.74. H. Xu et al. [30] believed that the birds found produced a noise filter sound and future testing should conduct tests to examine the effectiveness of this method. The approach used in the MICV and MIC-MFT is better than other selection approaches in terms of how well it classifies features

III EXISTING SYSTEM

We conducted a literature survey to find out what all work had been done in this field and what tools are needed to go ahead with the system. The survey gave us insights about the features required to classify the bird sounds and which models are needed for training to move further. The process of extraction of different features and selection of Mel Frequency Cepstral Coefficients to give input to the Support Vector Machine (SVM) as it gives a decent accuracy of about 89.64%. Extreme learning machines (ELM) are used to model emotional perception of audio and video features. ELM is used as an alternative to single layer feed forward networks for fast and accurate learning. A good open source tool for speech processing and music information retrieval is openSmile. This software provides easy extraction of audio features. The selection of features needs to be done properly as some of the features are such that including them reduces the accuracy of the model

IV PROPOSED SYSTEM

The (Female Feature MFCC) dataset was used in this work. It is an open-source dataset published on kaggle.com. This dataset created with the aim to predict female's emotions based on MFCCs values. With this setup (58 values for each emotion) we were been able to get a good 94% accuracy on the female emotions. The mean for a data set is termed as the arithmetic mean. In this paper, the mean of MFCCs is taken to reduce the huge set of values that are obtained from MFCC [33]. When 'x'= {x1, x2, x3, ,xn} represents MFCC values and total number of MFCC is n then the arithmetic mean is taken as

Type of trait in particular bird species expansion is one of the first things that needs to be done before bird classification can begin. certain species, the more basic types of bird songs are termed as note-like and song elements; in others, notes are the most simplistic. A regular string of words or phrases that are linked together with one following another is called a word sequence. The occurrence of one of the same series of one or more words or phrases at one time constitutes a motif or expression. the theory was developed based on this information, and prosody was employed to generate a catalogue of bird songs Since musical signals have the form (specific patterns of pronunciation, energy, and length), the prosody of a natural sound would be crucial to capture. Additionally, MFCCs can be used for the perception of emotion.

The procedure is completed, and then the several different classifications are considered to find out which approach is more commonly used these subsections will briefly explain the working concepts of the data classifiers that were used to build/test the training and test sets.

V IMPLEMENTATION

The first step is the process of collecting the bird sound data. Then, the collected data should be pre-processed to increase the efficiency of the playback. The next processes are features extraction and classification operations of sound patterns using machine learning algorithms based on derived features

Data Collection :

The (Female Feature MFCC) dataset was used in this work. It is an open-source dataset published on kaggle.com. This dataset created with the aim to predict female's emotions based on MFCCs values. With this setup (58 values for each emotion) we were been able to get a good 94% accuracy on the female emotions. The mean for a data set is termed as the arithmetic mean [32]. In this paper, the mean of MFCCs is taken to reduce the huge set of values that are obtained from MFCC [33]. When 'x'= {x1, x2, x3, ,xn} represents MFCC values and total number of MFCC is n then the arithmetic mean is taken

Pre-Processing :

A natural recording has a lot of background noise in it, so there is no way to clean up the audio clips prior to use [34]. Also, unwanted sounds have been used in the original sound recordings. it is essential to eliminate or minimize the noise from the desired sound, particularly when you are trying to hear the call of a wild bird Preprocessing of the sound recordings is needed to ensure reasonable device efficiency [35]. noise reduction methods (like filtering) must be applied to the recorded signals in order to take out unnecessary noise components that can, in turn, enhance the sound quality of the resulting recordings A frequency-based pass band-pass filter (Butter) is used to get rid of unwanted background noise as it filters out a frequency spectrum and has a very flat

frequency response in the passband [36]. For records of many species to have been discovered, it's found that overlapping sounds can be heard [37]. Nevertheless, we will conclude that there is only one dominant species of bird species this time around which is making the sounds from the tape [38]. The other sections of the audio signal, except the required ones, must be removed from the signal. To receive a known-requirements query that yields focused results, or tuned results, unbind the query so that it can return a single type of signal that has the results. in the way that the mean pitch is calculated using the harmonic continuum analysis [39]. From this, the preceding material, the Butterworth filter is manually calibrated, and a spectrogram is used to ensure that the waveforms have been recovered [40]. While the accuracy is increased, a downside of this approach is that other birds might still be making the same signal, which is when used alongside it [41]. But in this case, bird species have frequencies that are close to one another and appear to possess identical characteristics separating the various recordings of birds' calls to increase the success of singlelabel methods results would be very impractical so doing so would only make the success of the results of each form of detection less likely [42]. The instances that unable to match one another were taken as a consequence of the sheer volume of alternative recordings in our records. If we expand the analysis to two bands (a range of 30 and 15 Hz), we can see the results of using a Butterworth Hummingbird Recording in the audio archive [43]. Expanded on the last graph are 2a and 2b and 2c, the music recording of Anna's hummingbird song in our database exhibit the recorded and continuous waveforms [44]. Seen in Fig. 2d, which does not have any added amplitude enhancement, a boost the low frequencies on both sides of the cut off from about 5000 Hz. A suppressed noise waveform (as seen in Fig. 2b) results in a filtered output

(the opposite of the filtered output in which can be seen in Fig. 2a) however, as seen in Fig. 2a, there may be no meaningful detail in the initial signal while, as seen in the other Fig. 2b, may have useful features

Feature Extraction :

Type of trait in particular bird species expansion is one of the first things that needs to be done before bird classification can begin [46][47][48][49]. In certain species, the more basic types of bird songs are termed as note-like and song elements; in others, notes are the most simplistic [50]. A regular string of words or phrases that are linked together with one following another is called a word sequence [51]. The occurrence of one of the same series of one or more words or phrases at one time constitutes a motif or expression. the theory was developed based on this information, and prosody was employed to generate a catalogue of bird songs Since musical signals have the form (specific patterns of pronunciation, energy,

and length), the prosody of a natural sound would be crucial to capture [52]. Additionally, MFCCs can be used for the perception of emotion.

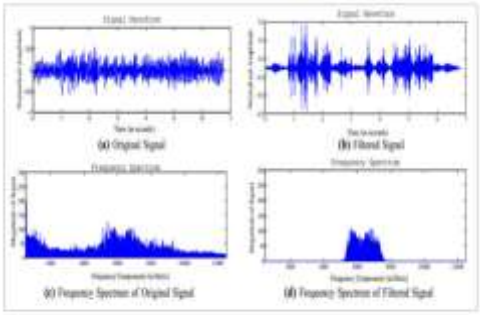

Fig. 2. Effect of butterworth filter on Anna's hummingbird song recording

Classification Methods :

K-Nearest Neighbor(KNN)

o K-Nearest Neighbour is one of the simplest Machine Learning algorithms based on Supervised Learning technique.

o K-NN algorithm assumes the similarity between the new case/data and available cases and put the new case into the category that is most similar to the available categories.

o K-NN algorithm stores all the available data and classifies a new data point based on the similarity. This means when new data appears then it can be easily classified into a well suite category by using K- NN algorithm.

o K-NN algorithm can be used for Regression as well as for Classification but mostly it is used for the Classification problems.

o K-NN is a non-parametric algorithm, which means it does not make any assumption on underlying data.

o It is also called a lazy learner algorithm because it does not learn from the training set immediately instead it stores the dataset and at the time of classification, it performs an action on the dataset.

o KNN algorithm at the training phase just stores the dataset and when it gets new data, then it classifies that data into a category that is much similar to the new data.

o Example: Suppose, we have an image of a creature that looks similar to cat and dog, but we want to know either it is a cat or dog. So for this identification, we can use the KNN algorithm, as it works on a similarity measure. Our KNN model will find the similar features of the new data set to the cats and dogs images and based on the most similar features it will put it in either cat or dog category.

Random Forest Algorithm

Random Forest is a popular machine learning algorithm that belongs to the supervised learning technique. It can be used for both Classification and Regression problems in ML. It is based on the concept of ensemble learning, which is a process of combining multiple classifiers to

solve a complex problem and to improve the performance of the model.

As the name suggests, "Random Forest is a classifier that contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset." Instead of relying on one decision tree, the random forest takes the prediction from each tree and based on the majority votes of predictions, and it predicts the final output.

The greater number of trees in the forest leads to higher accuracy and prevents the problem of over fitting.

MLP :

The multilayer perceptron MLP is one of the neural network types, utilizes a supervised learning technique called backpropagation for training. Supervised multilayer perceptron (MLP), was used as a classification of birds sounds in this paper. The neural networks model has been used due to its ability to compensate discrepancies in the data. This is one way to deal with the individual and regional variability of bird

Naïve Bayes :

Naïve Bayes algorithm is a supervised learning algorithm, which is based on Bayes theorem and used for solving classification problems.

It is mainly used in text classification that includes a high-dimensional training dataset.

Naïve Bayes Classifier is one of the simple and most effective Classification algorithms which helps in building the fast machine learning models that can make quick predictions.

It is a probabilistic classifier, which means it predicts on the basis of the probability of an object.

Some popular examples of Naïve Bayes Algorithm are spam filtration, Sentimental analysis, and classifying articles.

VI RESULT ANALYSIS

According to the methodology that we previously defined, we have implemented machine learning algorithms (Naïve Bayes, J4.8 and Multilayer Perceptron) on a data set MFCC) Using program WEKA, we obtained different results as shown in the following table:- The difference in accuracy between these algorithms is significant as shown in the above table. Naive Bayes algorithm's accuracy (47.45%) is less accurate than the rest and has less time (1.19 seconds), while the MLP algorithm is more accurate (74.68%) and time consuming (1838.19 seconds), while J4.8 has the highest accuracy (78.40%) and is the best. Accuracy and elapsed time are (39.4 seconds). ; Article no.AJRCOS.68530 Measure: Precision and Recall are combined to form the harmonic mean.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
IJARST
ISSN: 2457-0362

To put it another way, it's a mathematical method for evaluating a system's accuracy by taking into account both (5) RESULTS AND DISCUSSION According to the methodology that we previously defined, we have implemented machine learning algorithms (Naïve Bayes, J4.8 and Multilayer Perceptron) on a data set (Female Feature MFCC) Using program WEKA, we obtained different results as shown in the The difference in accuracy between these algorithms is significant as shown in the above table. Naive Bayes algorithm's accuracy (47.45%) is less accurate than the rest and has less time (1.19 seconds), while the MLP algorithm is more accurate (74.68%) and timeconsuming (1838.19 seconds), while J4.8 has the highest accuracy (78.40%) and is racy

Table 1. Result of all evaluation metrics by using Naïve Bayes

Class	TP Rate	FP Rate	Precision	Recall	F-Measure
fear	0.450	0.034	0.711	0.450	0.551
angry	0.431	0.041	0.659	0.431	0.521
disgust	0.637	0.230	0.339	0.637	0.442
neutral	0.420	0.012	0.852	0.420	0.563
sad	0.408	0.093	0.449	0.408	0.428
surprise	0.783	0.074	0.466	0.783	0.568
happy	0.313	0.085	0.406	0.313	0.354
calm	0.935	0.041	0.241	0.935	0.384

Table 2. Result of all evaluation metrics by using J4.8 algorithm

Class	TP Rate	FP Rate	Precision	Recall	F-Measure
fear	0.780	0.044	0.767	0.780	0.774
angry	0.828	0.034	0.816	0.828	0.822
disgust	0.737	0.050	0.732	0.737	0.734
neutral	0.780	0.034	0.784	0.780	0.782
sad	0.780	0.039	0.786	0.780	0.783
surprise	0.893	0.038	0.899	0.893	0.896
happy	0.751	0.043	0.765	0.751	0.758
calm	0.766	0.033	0.808	0.766	0.787

Table 3. Result of all evaluation metrics by using MLP

Class	TP Rate	FP Rate	Precision	Recall	F-Measure
fear	0.707	0.048	0.731	0.707	0.719
angry	0.834	0.039	0.811	0.834	0.823
disgust	0.651	0.052	0.699	0.651	0.674
neutral	0.740	0.054	0.683	0.740	0.710
sad	0.745	0.051	0.728	0.745	0.736
surprise	0.939	0.005	0.941	0.939	0.940
happy	0.715	0.050	0.728	0.715	0.721
calm	0.765	0.002	0.825	0.765	0.794

Table 4. Classification accuracy and time of the proposed classifiers

Classifier	Classification accuracy (%)	Time in seconds
Naïve Bayes	47.4504 %	1.19 seconds
J4.8	78.4008 %	39.4 seconds
MLP	74.681 %	1838.19 seconds

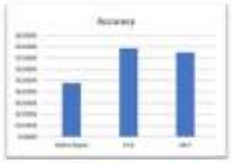

Fig. 6. Accuracy of the proposed algorithms

7

VII CONCLUSION

Machine learning algorithms have been used to classify and identify bird sound and sound emotion recognition. Bird species can be known by recording only the sound of the bird, which will make it easier for the system to manage. The system also provides species classification resources to allow automated species detection from observations that can teach a machine how to recognize whether or classify the species. In this work, Mel-frequency cepstral coefficient (MFCC)has been used and tested through different algorithms namely Naïve Bayes, J4.8 and Multilayer perceptron (MLP) in the classification of bird's species. The J4.8 algorithm shows the highest accuracy at 78.4008% and the time spent is 39.4 seconds.

REFERENCES

1. Charbuty B, Abdulazeez A. "Classification Based on Decision Tree Algorithm for Machine Learning," J. Appl. Sci. Technol. Trends. 2021;01:20–28, 2021. DOI: 10.38094/jastt20165.

2. Saleem SI, Abdulazeez AM, Orman Z. "A New Segmentation Framework for Arabic Handwritten Text Using Machine Learning Techniques; 2021, DOI: 10.32604/cmc.2021.016447.

3. Maulud D, Abdulazeez AM. "A Review on Linear Regression Comprehensive in Machine Learning," J. Appl. Sci. Technol. Trends. 2020;1(4):140–147. DOI: 10.38094/jastt1457.

4. Trappey AJC, Trappey CV, Wu JL, Wang JWC. "Intelligent compilation of patent summaries using machine learning and natural language processing techniques," Adv. Eng. Informatics. 2019- 220;43:101027. DOI: 10.1016/j.aei.2019.101027.

5. Abdulkareem NM, Abdulazeez AM. "Machine Learning Classification Based on Radom Forest Algorithm : A Review. 2021;128–142. DOI: 10.5281/zenodo.4471118.

6. Mahmood MR, Abdulazeez AM, Orman Z. "A New Hand Gesture Recognition System Using Artificial Neural Network," 2nd Int. Sci. Conf. ZAKHO. 2017;1–14.

7. Lasseck M. "Audio-based bird species identification with deep convolutional neural networks," CEUR Workshop Proc. 2018;2125.

8. Zeebaree DQ, Haron H, Abdulazeez AM, Zebari DA. "Machine learning and Region Growing for Breast Cancer Segmentation," 2019 Int. Conf. Adv. Sci. Eng. ICOASE 2019;88–93. DOI: 10.1109/ICOASE.2019.8723832

. 9. Ibrahim I, Abdulazeez A. "The Role of Machine Learning Algorithms for Diagnosing Diseases," J. Appl. Sci. Technol. Trends. 2021;2(01):10–19. DOI:

10.38094/jastt20179. 10. Piczak KJ. "Recognizing bird species in audio recordings using deep convolutional neural networks," CEUR Workshop Proc. 2016;1609:534–543.

11. Kahl S, Wilhelm-Stein T, Klinck H, Kowerko D, Eibl M. Recognizing birds from sound - The 2018 BirdCLEF baseline system," arXiv; 2018.

12. Abdulqader DM, Abdulazeez AM, Zeebaree DQ. "Machine learning supervised algorithms of gene selection: A review," Technol. Reports Kansai Univ. 2020;62(3);233–244.

13. Sprengel E, Jaggi M, Kilcher Y, Hofmann T. "Audio based bird species identification using deep learning techniques," CEUR Workshop Proc. 2016;1609:547–559.

14. Salim NOM, Abdulazeez AM. "Human Diseases Detection Based On Machine Learning Algorithms : A Review. 2021;102–113. DOI: 10.5281/zenodo.4467510.

15. Madhavi A, Pamnani R, Student PG. "Deep Learning Based Audio Classifier for Bird Species. 2018;S3(10):228–233.

16. Jaiprakash PS, Prakash CS, Desai M. "Image Splicing Forgery Detection Using DWT and Local Binary Pattern," Lect. Notes

Networks Syst. 2020;120(3):233– 242. DOI: 10.1007/978-981-15-3325-9_18.

17. Sadeeq H, Abdulazeez A, Kako N, Abrahim A. "A novel hybrid bird mating optimizer with differential evolution for engineering design optimization problems," Lect. Notes Data Eng. Commun. Technol. 2018;5:522–534. DOI: 10.1007/978-3-319-59427-9_55.

18. Gradišek A, et al. "How to recognize animal species based on sound-a case study on bumblebees, birds, and frogs. 2018;3–46

19. Clementino T, Colonna JG. "Using Triplet Loss for bird species recognition on BirdCLEF. 2020;22–25.

20. Mahmood MR, Abdulazeez AM. "A comparative study of a new hand recognition model based on line of features and other techniques," Lect. Notes Data Eng. Commun. Technol. 2018;5:420–432. DOI: 10.1007/978-3-319-59427-9_45.

21. V. Morfi, Y. Bas, H. Pamula, H. Glotin, and D. Stowell, "NIPS4Bplus: A richly annotated birdsong audio dataset," PeerJ Comput. Sci., vol. 2019, no. 10, pp. 1–12, 2019, doi: 10.7717/peerj-cs.223.

22. Pareta A, Taran S, Bajaj V, Sengur A. "Automatic Environment Sounds Classification Using Optimum Allocation Sampling," 2019 4th Int. Conf. Robot. Autom. Eng. ICRAE. 2019;69–73. DOI: 10.1109/ICRAE48301.2019.9043832.

23. Kyungdeuk KO, Park J, Han DK, Ko H. "Channel and frequency attention module for diverse animal sound classification," IEICE Trans. Inf. Syst., vol. E102D. 2019;12:2615–2618. DOI: 10.1587/transinf.2019EDL8128.

24. Lezhenin I, Bogach N, Pyshkin E. "Urban sound classification using long short-term memory neural network," Proc. Fed. Conf. Comput. Sci. Inf. Syst. FedCSIS 2019;18:57–60, 2019. DOI: 10.15439/2019F185.

25. Nanni L, Brahnam S, Lumini A, Maguolo G. Animal sound classification using dissimilarity spaces," Appl. Sci. 2020;10(23):1–18. DOI: 10.3390/app10238578.