Animal Identification Based on Voice Data using Classififcation Algorithm of Data Mining

Vivek Francis Pinto PG Student, Department of Computer Science and Engineering NMAM Institution of Technology, Nitte Karkala, India Vivekpinto07@gmail.com

Sampath Kini Assistant Professor, Department of Computer Science and Engineering NMAM Institution of Technology, Nitte Karkala, India Sampath@nitte.edu.in

Igneta Mcluren Dsouza PG Student, Department of Computer Networking and Engineering RV College of Engineering, Mysuru Road Bengaluru, India Mcluren@gmail.com

Abstract—Paper gives an idea about the implemented solution to classify the given voice as either Cat or Dog. Every voice is a signal which has different attributes called as features. Implementation has been done by considering three of such features as part of feature selection, discussed in the succession of the document. K-Nearest Neighbor algorithm has been used for performing classification and results have been tabulated for different percentage of test and training data. Work done helps to conclude that signal features can be used to classify the voice data.

Keywords-Voice Based Classification; Feature Selection; K-Nearest Neighbour; Supervised Learning;

I. INTRODUCTION

Nature consists of varieties of animals with different quality of their voices. Each voice is a signal having varieties of attributes in it, called as features [1]. Using the feature collection a voice signal can be categorized into different categories as needed. On the technology front, Data mining is the process of determining the patterns from the collection of data [2]. Mining consists of Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation Model and Deployment [2].

Fig 1 represents the six phases in the Data Mining Process. Business understanding helps to set the objectives for the business, goal for mining, validate current situation of business and overall project plan building. Data Understanding comprises various sources of data, quality check and data description. Data Preparation is the step upon data understanding which performs data cleaning and data transformation. Model building is the phase in which initial analysis is performed using graphs, charts or any viewing technology so that data can be grouped and data is modelled as test and training data in this step. Evaluation and Testing is the phase which actually puts the data that need to be mined into different tests and come up with the evaluations to build the confusion matrix that helps in assessing the test cases. Deployment is the phase in which the evaluated model is used a full pledged application that helps in business operations.

Data mining task generally divided into Predictive and Descriptive task [3]. In this paper, methods used for performing the voice classification, type of Predictive task is explained and phases of data collection, feature extraction, KNN algorithm usage and evaluation and testing is explained. K-Nearest Neighbor is a non-parametric method used for classification. The output of the KNN class is a label or class out of the defined label or class. For example, our implementation considers on Cat and Dog voice samples for the experiment, so the output is either Cat or Dog only.



Figure 1. Data Mining Phases

II. DATA COLLECTION, PREPROCESSING, FEATURE EXTRACTION

Data Collection: In this phase data from different sources categorized as online and manual collection was collected. By online collection we mean is that the audio recordings of Cat and Dog voices where collected. Mobile phone recorder and audio recorder was used to capture the voice of Cat and Dog at home, parking slots and streets. Voice samples have been downloaded from sites over the internet as part of online data collection phase.

Preprocessing: To keep the voice samples as .wav type files, files of non .wav type converted the files to .wav files as part of the preprocessing phase. TABLE 1 shows few of the samples of the data collected and preprocessed. By preprocessing it means the few of the files which were of mp3 format have been converted to .wav file type and few videos have been truncated to reduce the file size.

| Sl.No | File Name | File Size(KB) | File Type |
|-------|-----------|---------------|--------------------|
| 1 | Catl | 51.6 | Wave(Preprocessed) |
| 2 | Cat2 | 34.1 | Wave |
| 3 | Cat3 | 57.6 | Wave |
| 4 | Cat4 | 55.7 | Wave(Preprocessed) |
| 5 | Dog1 | 56.5 | Wave |
| 6 | Dog2 | 50.1 | Wave(preprocessed) |
| 7 | Dog3 | 70.2 | Wave |
| 8 | Dog4 | 86.2 | Wave |

TABLE 1: SAMPLE VOICE RECORD FILES OF CAT AND DOG

Feature extraction: As stated earlier in the document, features helps to distinguish the voice of different samples under consideration. Three features have been considered for the experiment. The Process that's followed for extracting the values of different attributes of the signal is called as feature extraction [4].Frequency defined as number of oscillations of the wave per second [5].Zero Crossing is defined as the number of times the waves crosses the 'zero' line [6] and Power which is the square of the frequency of the signal per frame [7]. TABLE 2 shows the samples of the extracted values for the collected data set of Cat and Dog Voices. Further in this step the class labels for each record of features is also attached which makes it as a supervised learning. Observe the TABLE 2 under column labeled as 'Class Label' that tell the class of the record if it is Cat or Dog voice.1 Represents Cat and 2 Represents Dog.

| Zero Crossing | Frequency | Power | Class Label |
|------------------|-----------|----------|----------------|
| 846 | 3.19E+02 | 4.96E+01 | 1 |
| 1242 | 2.88E+02 | 4.06E+01 | 1 |
| 754 | 9.06E+01 | 4.00E+00 | 1 |
| 716 | 1.33E+02 | 8.65E+00 | 1 |
| 754 | 9.06E+01 | 4.00E+00 | 2 |
| 716 | 1.33E+02 | 8.65E+00 | 2 |
| 647 | 1.33E+02 | 8.65E+00 | 2 |
| 735 | 1.37E+02 | 9.15E+00 | 2 |

TABLE 2: SAMPLE OF THE FEATURES OF VOICE FILES WITH CLASS LABEL

III. MODELLING , CLASSIFICATION AND EVALUATION

Upon the process of feature extraction features are represented using a 3D graph which helps to do a classification to the possible extent just by seeing the graph generated with the feature values of Frequency, Zero crossing and Power of each frame for different voice samples. In the graph Red indicates the Cat Label and Blue indicates the Dog Label. Fig 2 shows the modelling of the extracted features of the voice samples considered for the experimentation and we can draw a line that helps us to imagine two classes one per side of line.

In order to do the classification.K-Nearest Neighbour algorithm is used which is used to calculate the distance between the set of point.If the distance is zero or close to zero, it indicated that the two points (here feature values of voice sample ,frequancy ,zero crossing and power) are similar in nature.Classification is the heart of the implementation that has been performed, weightage lies on classifying the test records based on the training records. However, the available training data has been portioned so that, one partition is used for testing purpose. The data with actual label is partitioned into 2 sets called the training and the test sets. A classification model is then induced from the training set and its performance is evaluated on the test set.

Different combination of partitions has been considered for testing purpose. Data is divided in 3 different percentage of training record and testing records. First type partition consists of 50% of the record in training and 50% records in testing set, second type partition consists of 75% of records in training and 25% in test set and the last partition consists of 25% of records in training set and 75% in testing set.

upon having values of training data comes to a conclusion that the given test record possibly is classified as corresponding label for which it finds the minimum in the sorted test records upon calculating the distance from each training records using the equation 1.

Distance = root (square(testFeatureX-trainingFeatureX) +square(testFeatureY-trainingFeatureY) +square (testFeatureZ-trainingFeatureZ)) (1)

Where testFeatureX,Y,Z are the Frequency,Zero Crossing and Power feature of the test record and trainingFeatureX,Y,Z are the Frequency,Zero Crossing and Power feature of Training records.

The algorithm is KNN in which K is an integer that needs to be selected to set the limit for sorting of distance and selecting the minmum distance among K calculated distance from a test point(X,Y,Z) attribute to training



Figure 2 : Modelling of the feature values using 3D Graph TABLE 3: SAMPLE OF THE CLASSIFICATION PERFORMED

| Sl No | Zero Crossing | Frequency | Power | Actual Label | Predicted Label |
|-------|---------------|-------------------|--------------------|--------------|-----------------|
| 1 | 262 | 0.257774140752864 | 3.244500902722e-05 | 1 | 1 |
| 2 | 0 | 0.257774140752864 | 3.244507209722e-05 | 1 | 1 |
| 3 | 388 | 71.4074857970369 | 2.48976026750685 | 1 | 1 |
| 4 | 368 | 35.4405703464409 | 0.613297864492689 | 1 | 1 |
| 5 | 317 | 30.3525676729420 | 0.449842951338145 | 1 | 2 |
| 6 | 258 | 0.175448368051688 | 1.503033684179e-05 | 2 | 1 |
| 7 | 576 | 141.977163461538 | 9.84253659403536 | 2 | 2 |
| 8 | 1012 | 143.239182692308 | 10.0182927042775 | 2 | 1 |
| 9 | 462 | 140.084134615385 | 9.58181873581113 | 2 | 2 |
| 10 | 435 | 135.036057692308 | 8.90368011576184 | 2 | 2 |

point (X,Y,Z).Experimental Result section lists the different K values that are considered and percentage of the accuracy of the implemented model for different percentage of the test and training records. TABLE 3 show some of the classification done, the fourth column is the actual label and 5th column is the

predicted label upon classification using the learning model deduced by the training records. To help the reader to understand the table, the second row means that the record having those values is a Cat voice (Observed as part of training step) and it was classified as Cat Voice in classification also.

IV. EXPERIMENTAL RESULST

As stated earlier, the implemented algorithm has been tested using the different scenarios, few samples scenarios have been listed here. Once the classification model is deduced using the training records, the classification model is put into use and test records are predicted using the model. However, the process doesn't guarantee that the classification is 100% correct. So different number of training and testing records are considered as explained earlier and the classification process is performed and results are captured.

The experiment helps to conclude that I classification model can be built using the KNN model for performing voice based classification. For the collected data, the experiment showed accuracy of 86.36% for the case when records where spilt 50% as training records and 50% as test records. The experiment shows that the good value for K varies for different % of the records and in the experiment performed ,86.36 accuracy was achieved for the value K= 5.Finally the model was extended to take a voice sample as input and classify the file as either Cat or Dog.

| Value of K | Partition Percentage | Accuracy (in %) |
|------------|-----------------------------------|-----------------|
| 3 | | 51.1363 |
| 5 | 50% of records as training record | 86.3636 |
| 7 | and 50% as test records | 72.7273 |
| 9 | | 71.5909 |
| 3 | | 51.1111 |
| 5 | 75% of records as training record | 51.1111 |
| 7 | and 25% as test records | 51.1111 |
| 9 | | 51.1111 |
| 3 | | 40.6015 |
| 5 | | 39.8496 |
| 7 | 25% of records as training | 33.8346 |
| 9 | - records and 75% as test records | 11.2782 |

TABLE 3: SAMPLE OF THE CLASSIFICATION PERFORMED USING KNN

CONCLUSION

Data mining is one of potential technology in the current world which is generating lot of data ,more or less upon every single activity ,the data is getting generated and being collected. Implementation work has helped to get the understanding of the potential of Data mining and to come up with an application that would classify the given voice as either Cat or Dog. Further experimental results shows that tweaking between the values selection for K in KNN algorithm also influences the accuracy of the classification. Further the quality and size of the training set also influences the accuracy of the classification.

REFERENCES

- [1] Graf, S., Herbig, T., Buck, M. et al. EURASIP J. Adv. Signal Process. (2015) 2015: 91. https://doi.org/10.1186/s13634-015-0277-z
- [2] David L. Olson, Dursun Delen. Advanced Data Mining Techniques. (2015). https://doi.org/10.1007/978-3-540-76917-0
- [3] Sonali Guglani, Sunaina Bagga, Ankit Goyal, "Desciption about predictive and descriptive data mining". International journal of advanced and innovative research(IJAIR),2012, ISSN: 2278-7844,Page(s):303-306
- [4] M. Pechenizkiy, S. Puuronen, A. Tsymbal," Feature extraction for classification in the data mining process". International Journal "Information Theories & Applications",2003, ISSN 1310-0513 ,Vol 10,Page(s):271-278
- [5] A David, Handbook of Condition Monitoring. Chapman and Hall, Uk: University of Wales-Cardiff, 1998
- [6] Domenicantonio Grillo, et.al, "An Efficient Extension of the Zero-Crossing Technique to Measure Frequency of Noisy Signals", 978-1-4577-1772-7/12/\$26.00 ©2012 IEEE.
- [7] Shilpa.R,Savitha Shetty K L, P. S. Puttaswamy, "Analysis of Signals Affecting the Quality of Power", Journal of Signal Processing and Wireless Technologies(JSPWT),2017,Vol 1,Page(s)10-13
- [8] P.Daponte D.Grimaldi A.Molinaro Ya.D.Sergeyev," Fast detection of the first zero-crossing in a measurement signal set", ScienceDirect, Elsevier, 1996, Vol 19, Issue 1: Page(s): 29-39
- [9] Jiawei Han and Micheline Kamber (2006), Data Mining Concepts and Techniques, published by Morgan Kauffman, 2nd ed.
- [10] Karel Dejaeger, Wouter Verbeke, David Martens, and Bart Baesens," Data Mining Techniques for Software Effort Estimation: A Comparative Study", IEEE Transactions On Software Engineering, Vol. 38, No. 2, March/April 2012