Jason Kozodoy
Conor Van Achte
Neil Prakasam
Assignment 4 Report (Option 1)
06/06/17

1. How to run the code, how to use the system, functionalities of each program file

Double click the jar file to run the program. Don't take the jar file out of the Assignment4 folder. Running the Jar file will create a data file that the system will use when run. The program makes use of the GUI used for the previous assignment and has 40 sound files and four different data mining techniques available for the user to click on. The user can view the music files on the first page and audio files on the second page. The user can also click on the music or audio file and listen to its sound. We use four methods of analyzing the audio files data. The four methods (classifiers) are J48, ZeroR, Naive Bayes, and SMO (Sequential Minimal Optimization). J48 uses a decision tree and ZeroR uses a 0-R classifier and predicts the mode and mean. The Naive Bayes creates precision values that is based on training data. SMO uses the sequential minimal optimization algorithm with support vectors.

2. List and briefly introduce libraries/tools/techniques you used in your development

The tool we used for this assignment is called WEKA. We extract the features and data of the audio files and write this data to the .arff file. Once the data is written to the .arff file then the WEKA data mining program can begin its functionality. We also use the java API musicg to categorize the data within the audio files.

3. List features and their values for the audio files and show what files are used as training data, what are testing.

From the audio files we are getting features like byte rate, channels, frames per second, maximum amplitude, minimum amplitude, and other features that are included in the output below. We put all this data in one file and then use WEKA'S training methods.

4. Show comparison between model output and ground truth label (as discussed in 2.4) and display the precision and recall value of your model on the testing data.

J48:

J48 pruned tree

Chunk Size <= 128034: MUSIC (8.0)

```
Chunk Size > 128034
|   Chunk Size <= 128036
|   |   Maximum Amplitude <= 2749: SPEECH (6.0)
|   |   Maximum Amplitude > 2749
|   |   |   Mean <= 0.000064: SPEECH (13.0/5.0)
|   |   |   Mean > 0.000064: MUSIC (3.0)
|   Chunk Size > 128036
|   |   Mean <= -0.0003: SPEECH (5.0)
|   |   Mean > -0.0003: MUSIC (5.0/1.0)

Number of Leaves  :         6

Size of the tree :    11


Time taken to build model: 0.01 seconds
Time taken to test model on training data: 0 seconds

=== Error on training data ===

Correctly Classified Instances          34               85      %
Incorrectly Classified Instances         6               15      %
Kappa statistic                          0.7
Mean absolute error                      0.1938
Root mean squared error                  0.3113
Relative absolute error                 38.7692 %
Root relative squared error             62.2649 %
Total Number of Instances               40


=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area  Class
                 0.75      0.05       0.938     0.75      0.833      0.933    MUSIC
                 0.95      0.25       0.792     0.95      0.864      0.933    SPEECH
Weighted Avg.    0.85      0.15       0.865     0.85      0.848      0.932


=== Confusion Matrix ===

  a  b   <-- classified as
 15  5 |  a = MUSIC
  1 19 |  b = SPEECH
```

=== Stratified cross-validation ===

Correctly Classified Instances	25	62.5 %
Incorrectly Classified Instances	15	37.5 %
Kappa statistic	0.25	
Mean absolute error	0.4047	
Root mean squared error	0.5503	
Relative absolute error	80.9405 %	
Root relative squared error	110.0605 %	
Total Number of Instances	40	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.55	0.3	0.647	0.55	0.595	0.631	MUSIC
	0.7	0.45	0.609	0.7	0.651	0.631	SPEECH
Weighted Avg.	0.625	0.375	0.628	0.625	0.623	0.631	

=== Confusion Matrix ===

a	b	<-- classified as
11	9	a = MUSIC
6	14	b = SPEECH

SMO:

Options: -C 1.0 -L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1

SMO

Kernel used:
Linear Kernel: K(x,y) = <x,y>

Classifier for classes: MUSIC, SPEECH

BinarySMO

Machine linear: showing attribute weights, not support vectors.

```
        1.0441 * (normalized) Chunk Size
 +      -1.2636 * (normalized) Maximum Amplitude
 +      -0.2505 * (normalized) Minimum Amplitutde
 +       0.2132 * (normalized) Mean
 +       0.3153 * (normalized) Zero Crossing Rate
 +       0.4379 * (normalized) Standard Deviation
 +      -1.3181 * (normalized) Spectral Centroid
 +       0.826

Number of kernel evaluations: 306 (69.308% cached)



Time taken to build model: 0.04 seconds
Time taken to test model on training data: 0 seconds

=== Error on training data ===

Correctly Classified Instances          28               70      %
Incorrectly Classified Instances        12               30      %
Kappa statistic                          0.4
Mean absolute error                      0.3
Root mean squared error                  0.5477
Relative absolute error                 60      %
Root relative squared error            109.5445 %
Total Number of Instances               40


=== Confusion Matrix ===

  a  b   <-- classified as
 14  6 |  a = MUSIC
  6 14 |  b = SPEECH



=== Stratified cross-validation ===

Correctly Classified Instances          24               60      %
Incorrectly Classified Instances        16               40      %
Kappa statistic                          0.2
Mean absolute error                      0.4
```

Root mean squared error 0.6325
Relative absolute error 80 %
Root relative squared error 126.4911 %
Total Number of Instances 40

=== Confusion Matrix ===

 a b <-- classified as
14 6 | a = MUSIC
10 10 | b = SPEECH

Zero R:

ZeroR predicts class value: MUSIC

Time taken to build model: 0 seconds
Time taken to test model on training data: 0 seconds

=== Error on training data ===

Correctly Classified Instances 20 50 %
Incorrectly Classified Instances 20 50 %
Kappa statistic 0
Mean absolute error 0.5
Root mean squared error 0.5
Relative absolute error 100 %
Root relative squared error 100 %
Total Number of Instances 40

=== Confusion Matrix ===

 a b <-- classified as
20 0 | a = MUSIC
20 0 | b = SPEECH

=== Stratified cross-validation ===

Correctly Classified Instances 20 50 %
Incorrectly Classified Instances 20 50 %

Kappa statistic 0
Mean absolute error 0.5
Root mean squared error 0.5
Relative absolute error 100 %
Root relative squared error 100 %
Total Number of Instances 40

=== Confusion Matrix ===

 a b <-- classified as
20 0 | a = MUSIC
20 0 | b = SPEECH

Naive Bayes:

Naive Bayes Classifier

Attribute	Class MUSIC (0.5)	SPEECH (0.5)
Audio Format		
mean	1	1
std. dev.	0.0017	0.0017
weight sum	20	20
precision	0.01	0.01
Sample Rate		
mean	16000	16000
std. dev.	0.0017	0.0017
weight sum	20	20
precision	0.01	0.01
Bits Per Sample		
mean	16	16
std. dev.	0.0017	0.0017
weight sum	20	20
precision	0.01	0.01
Block Align		
mean	2	2
std. dev.	0.0017	0.0017

weight sum	20	20
precision	0.01	0.01

Byte Rate

mean	32000	32000
std. dev.	0.0017	0.0017
weight sum	20	20
precision	0.01	0.01

Channels

mean	1	1
std. dev.	0.0017	0.0017
weight sum	20	20
precision	0.01	0.01

Chunk Size

mean	128076	128089
std. dev.	52	59.5735
weight sum	20	20
precision	65	65

FFT Sample Size

mean	1024	1024
std. dev.	0.0017	0.0017
weight sum	20	20
precision	0.01	0.01

Frames Per Second

mean	15	15
std. dev.	0.0017	0.0017
weight sum	20	20
precision	0.01	0.01

Number of Frames

mean	62	62
std. dev.	0.0017	0.0017
weight sum	20	20
precision	0.01	0.01

Number of Frequency Units

mean	512	512
std. dev.	0.0017	0.0017
weight sum	20	20

precision	0.01	0.01
Unit Frequency		
mean	15.62	15.62
std. dev.	0.0017	0.0017
weight sum	20	20
precision	0.01	0.01
Time Step		
mean	0.1	0.1
std. dev.	0.0017	0.0017
weight sum	20	20
precision	0.01	0.01
Maximum Amplitude		
mean	36255.6154	23181.1346
std. dev.	17116.4732	19820.3669
weight sum	20	20
precision	1604.2308	1604.2308
Minimum Amplitutde		
mean	25402.2282	18070.3897
std. dev.	19319.4585	16236.3217
weight sum	20	20
precision	1481.1795	1481.1795
Mean		
mean	0	0.0013
std. dev.	0.0004	0.0032
weight sum	20	20
precision	0.0003	0.0003
Zero Crossing Rate		
mean	7841.2115	8127.8795
std. dev.	3470.4242	2799.2282
weight sum	20	20
precision	337.2564	337.2564
Standard Deviation		
mean	0.0909	0.1389
std. dev.	0.061	0.1339
weight sum	20	20
precision	0.0103	0.0103

Spectral Centroid

mean	33210.5895	30841.2252
std. dev.	2305.6538	3222.4419
weight sum	20	20
precision	394.894	394.894

Time taken to build model: 0.01 seconds
Time taken to test model on training data: 0.01 seconds

=== Error on training data ===

Correctly Classified Instances	25	62.5 %
Incorrectly Classified Instances	15	37.5 %
Kappa statistic	0.25	
Mean absolute error	0.3649	
Root mean squared error	0.5464	
Relative absolute error	72.9809 %	
Root relative squared error	109.285 %	
Total Number of Instances	40	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.9	0.65	0.581	0.9	0.706	0.783	MUSIC
	0.35	0.1	0.778	0.35	0.483	0.783	SPEECH
Weighted Avg.	0.625	0.375	0.679	0.625	0.594	0.783	

=== Confusion Matrix ===

a	b	<-- classified as
18	2	a = MUSIC
13	7	b = SPEECH

=== Stratified cross-validation ===

Correctly Classified Instances	20	50 %

Incorrectly Classified Instances	20	50	%
Kappa statistic	0		
Mean absolute error	0.4693		
Root mean squared error	0.6256		
Relative absolute error	93.8554 %		
Root relative squared error	125.127 %		
Total Number of Instances	40		

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.7	0.7	0.5	0.7	0.583	0.555	MUSIC
	0.3	0.3	0.5	0.3	0.375	0.555	SPEECH
Weighted Avg.	0.5	0.5	0.5	0.5	0.479	0.555	

=== Confusion Matrix ===

a	b	<-- classified as
14	6	a = MUSIC
14	6	b = SPEECH