Analysis

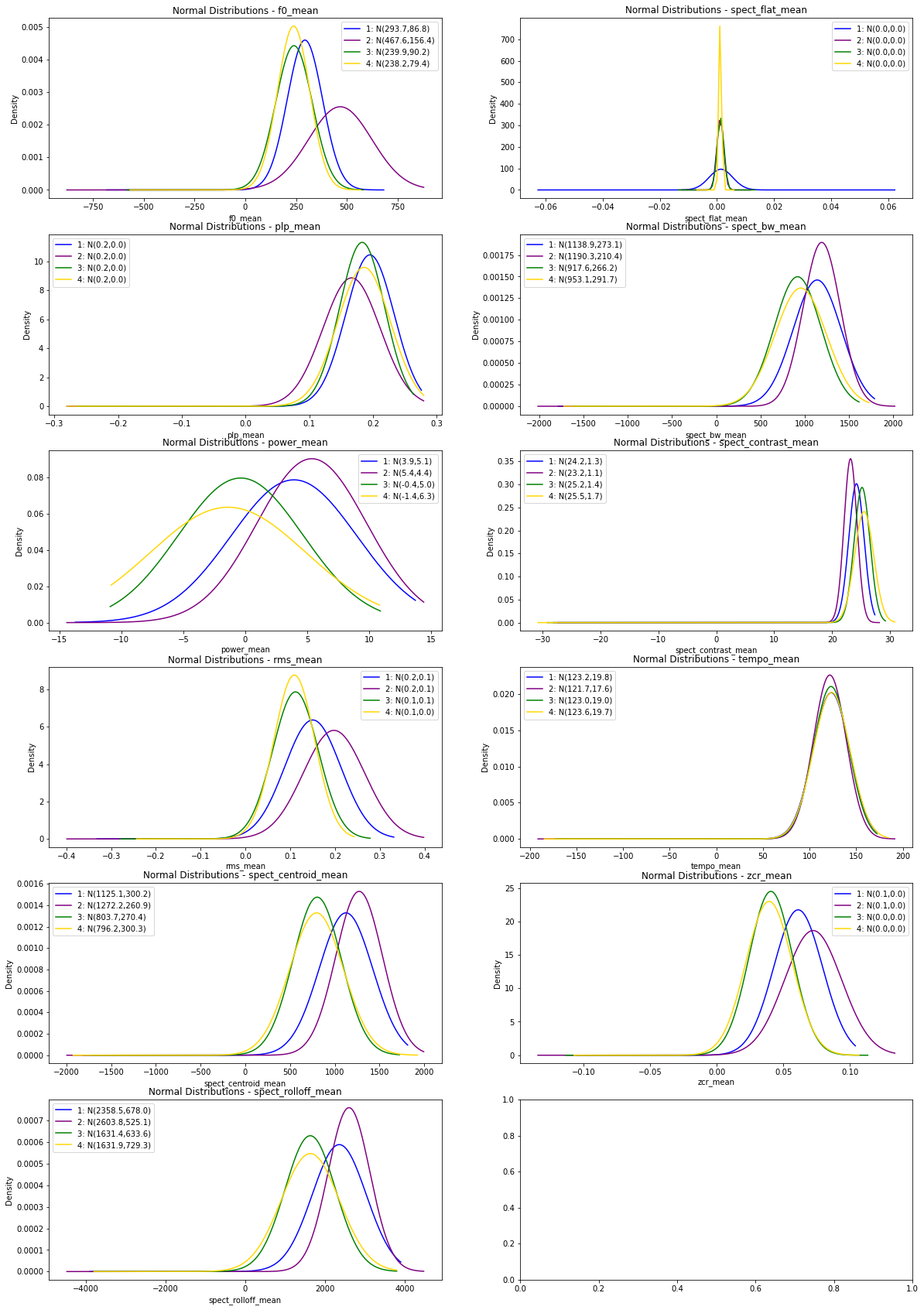
The dataset, MER Audio Traffic Data, contains sound samples classified into 4 quadrants that were adapted from Russell’s circumplex model of emotion. The dataset (N=900) has 225 samples of sound in each quadrant thereby making it a balanced dataset. The table below illustrates characterestics of several features of the sound samples in this dataset.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| N = 900, Q1=225, Q2= 225, Q3 = 225, Q4 = 225 | | | | | | |
| feature | class | mean | median | max | sd | iqr |
| f0 | Q1 | 169.229 - 682.105 | 132.501 - 806.396 | 1050 - 1050 | 76.135 - 362.434 | 24.743 - 817.885 |
| Q2 | 178.455 - 878.958 | 130.473 - 1050.0 | 1050 - 1050 | 107.8 - 440.867 | 27.794 - 919.527 |
| Q3 | 136.67 - 582.762 | 130.473 - 615.892 | 1050 - 1050 | 44.726 - 382.083 | 0.0 - 770.725 |
| Q4 | 151.714 - 567.331 | 130.473 - 486.91 | 1050 - 1050 | 48.355 - 403.464 | 0.0 - 854.414 |
| plp | Q1 | 0.097 - 0.277 | 0.0 - 0.033 | 1-Jan | 0.141 - 0.338 | 0.172 - 0.603 |
| Q2 | 0.044 - 0.28 | 0.0 - 0.013 | 1-Jan | 0.099 - 0.342 | 0.06 - 0.614 |
| Q3 | 0.098 - 0.264 | 0.0 - 0.04 | 1-Jan | 0.155 - 0.324 | 0.106 - 0.575 |
| Q4 | 0.076 - 0.28 | 0.0 - 0.016 | 1-Jan | 0.131 - 0.342 | 0.11 - 0.616 |
| power | Q1 | -14.589 - 13.717 | -6.189 - 21.679 | 22.041 - 39.348 | 18.159 - 33.365 | 6.484 - 69.119 |
| Q2 | -10.758 - 14.389 | 0.699 - 22.839 | 15.578 - 39.126 | 14.07 - 29.502 | 9.01 - 45.922 |
| Q3 | -14.938 - 10.896 | -9.284 - 17.833 | 16.806 - 38.551 | 19.03 - 31.429 | 8.361 - 71.451 |
| Q4 | -21.873 - 10.813 | -19.215 - 21.422 | 14.559 - 39.492 | 17.155 - 30.082 | 8.846 - 70.988 |
| rms | Q1 | 0.031 - 0.332 | 0.026 - 0.341 | 0.126 - 0.77 | 0.021 - 0.16 | 0.021 - 0.214 |
| Q2 | 0.022 - 0.399 | 0.021 - 0.43 | 0.058 - 0.689 | 0.008 - 0.15 | 0.009 - 0.246 |
| Q3 | 0.016 - 0.28 | 0.008 - 0.294 | 0.083 - 0.7 | 0.013 - 0.17 | 0.012 - 0.324 |
| Q4 | 0.017 - 0.243 | 0.017 - 0.235 | 0.048 - 0.75 | 0.009 - 0.142 | 0.011 - 0.203 |
| spect\_bw | Q1 | 393.436 - 1786.716 | 393.844 - 1843.467 | 964.397 - 3193.479 | 68.176 - 480.811 | 54.672 - 636.891 |
| Q2 | 515.302 - 2015.781 | 491.474 - 2096.548 | 1312.649 - 3136.144 | 61.459 - 407.248 | 62.439 - 760.684 |
| Q3 | 291.586 - 1611.415 | 267.225 - 1643.87 | 1039.565 - 3255.02 | 51.257 - 507.864 | 51.458 - 742.657 |
| Q4 | 196.51 - 1722.461 | 190.308 - 1942.784 | 1177.24 - 3385.334 | 55.389 - 493.28 | 48.233 - 738.594 |
| spect\_centroid | Q1 | 373.015 - 1814.521 | 368.615 - 1773.51 | 1079.09 - 7653.906 | 65.577 - 945.845 | 73.341 - 1674.926 |
| Q2 | 434.031 - 1994.109 | 379.542 - 1996.843 | 1561.807 - 7621.4 | 102.138 - 766.088 | 120.463 - 1277.371 |
| Q3 | 315.243 - 1722.115 | 271.812 - 1675.192 | 1363.449 - 7693.994 | 77.2 - 733.496 | 97.239 - 882.259 |
| Q4 | 186.474 - 1924.253 | 183.155 - 2221.394 | 1051.875 - 7635.677 | 36.68 - 866.793 | 39.372 - 1666.57 |
| spect\_contrast | Q1 | 19.869 - 27.383 | 16.0 - 26.021 | 41.904 - 82.556 | 7.527 - 19.834 | 7.091 - 24.035 |
| Q2 | 18.816 - 28.167 | 13.967 - 27.934 | 48.891 - 82.558 | 8.592 - 20.738 | 5.564 - 23.014 |
| Q3 | 20.57 - 29.201 | 15.717 - 28.66 | 43.077 - 82.946 | 7.571 - 19.859 | 6.909 - 21.894 |
| Q4 | 20.339 - 30.835 | 17.174 - 30.246 | 44.606 - 82.765 | 5.993 - 19.573 | 7.63 - 21.456 |
| spect\_flat | Q1 | 0.0 - 0.063 | 0.0 - 0.0 | 1.0 - 1.0 | 0.02 - 0.242 | 0.0 - 0.0 |
| Q2 | 0.0 - 0.012 | 0.0 - 0.0 | 1.0 - 1.0 | 0.02 - 0.109 | 0.0 - 0.0 |
| Q3 | 0.0 - 0.014 | 0.0 - 0.0 | 0.517 - 1.0 | 0.01 - 0.115 | 0.0 - 0.0 |
| Q4 | 0.001 - 0.007 | 0.0 - 0.0 | 1.0 - 1.0 | 0.021 - 0.082 | 0.0 - 0.0 |
| spect\_rolloff | Q1 | 660.864 - 3903.849 | 592.163 - 4080.542 | 2099.487 - 9420.776 | 161.71 - 1607.156 | 172.266 - 2640.509 |
| Q2 | 795.336 - 4479.259 | 742.895 - 4737.305 | 3445.312 - 8839.38 | 184.569 - 1262.966 | 204.565 - 2422.485 |
| Q3 | 500.531 - 3800.328 | 398.364 - 4037.476 | 3100.781 - 9087.012 | 189.785 - 1673.362 | 161.499 - 2842.383 |
| Q4 | 259.488 - 3796.873 | 258.398 - 4586.572 | 2487.085 - 9076.245 | 54.761 - 1512.542 | 43.066 - 2767.017 |
| tempo | Q1 | 89.103 - 172.266 | 89.103 - 172.266 | 89.103 - 172.266 | 0 - 0 | 0 - 0 |
| Q2 | 89.103 - 191.406 | 89.103 - 191.406 | 89.103 - 191.406 | 0 - 0 | 0 - 0 |
| Q3 | 87.593 - 172.266 | 87.593 - 172.266 | 87.593 - 172.266 | 0 - 0 | 0 - 0 |
| Q4 | 82.031 - 184.57 | 82.031 - 184.57 | 82.031 - 184.57 | 0 - 0 | 0 - 0 |
| zcr | Q1 | 0.017 - 0.105 | 0.011 - 0.101 | 0.081 - 0.444 | 0.01 - 0.096 | 0.012 - 0.161 |
| Q2 | 0.023 - 0.134 | 0.017 - 0.13 | 0.077 - 0.443 | 0.01 - 0.085 | 0.014 - 0.127 |
| Q3 | 0.015 - 0.114 | 0.012 - 0.114 | 0.047 - 0.555 | 0.004 - 0.064 | 0.005 - 0.078 |
| Q4 | 0.01 - 0.107 | 0.007 - 0.098 | 0.047 - 0.438 | 0.004 - 0.068 | 0.006 - 0.101 |

Columns to be ignored due to constant values - 'f0\_max', 'plp\_max', 'spect\_flat\_median', 'spect\_flat\_max', 'spect\_flat\_iqr', 'tempo\_sd', 'tempo\_iqr', 'f0\_max', 'plp\_max', 'spect\_flat\_median', 'spect\_flat\_max', 'spect\_flat\_iqr', 'tempo\_sd', 'tempo\_iqr', 'f0\_max', 'plp\_max', 'spect\_flat\_median', 'spect\_flat\_iqr', 'tempo\_sd', 'tempo\_iqr', 'f0\_max', 'plp\_max', 'spect\_flat\_median', 'spect\_flat\_max', 'spect\_flat\_iqr', 'tempo\_sd', 'tempo\_iqr'

The feature set is non-normal and will need standardization and/or normalization before modelling.

The plot below illustrates the normal distribution of mean values of key features across each class.



The k-means clustering also indicates a class overlap with the model indicating 3 classes in the dataset instead of 4, as illustrates in the figures below (expand this)

|  |  |
| --- | --- |
|  |  |

**Experiment 1 – Emotion model with raw and Hilbert transformed dataset with 4 classes**

* Data set DS\_1 – MER Audio Traffic Dataset
* Generate 2 datasets
  + Features from raw signal
  + Features from Hilbert transformed signal
* Extract 5 point summaries of following features associated with both datasets
  + Fundamental Frequency (F0)
  + RMS
  + Spectral Centroid
  + Spectral RollOff
  + Spectral Flatness
  + Spectral Bandwidth
  + Spectral Contrast
  + Zero Crossing Rate
  + Tempo
  + Predominant Local Pulse (PLP)
  + Power
  + 20 MFCC’s
  + 64 Mel Frequencies
  + Loudness
  + Chromagram
* The 5 point summaries considered are:
  + Mean
  + SD
  + IQR
  + Min and
  + Max
* Dependent variable – Russel’s Emotion quadrant assignment – Q1,Q2,Q3 and Q4
* Independent variables
  + Set 1: Raw features
  + Set 2: Hilbert Transformed features
* Modelling / Classification methods
  + AdaBoost Classifier
  + GradientBoosting Classifier
  + RandomForest Classifier
  + Support Vector Classifier
    - Kernel = ‘rbf’
    - Kernel =’poly’
    - Kernel = ‘linear’
  + Voting Classifier (Hard voting)
    - AdaBoost Classifier
    - GradientBoosting Classifier
    - RandomForest Classifier
    - Support Vector Classifier
  + Stacking Classifier
    - AdaBoost Classifier
    - GradientBoosting Classifier
    - RandomForest Classifier
    - Support Vector Classifier

Step 1: Consider only mean features

Visualization

Key features distribution (mean values) – F0, RMS, Loudness, Chromagram

Results

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| DataSet | Model | Features | Mean Accuracy & Range | F1 Score | AUC |
| Raw | AdaBoost | Mean of all features | 56.9 (54.1 – 61.5) | 55.9 (52.6 – 61.5) | 0.732 (71.1 – 79.4) |
|  |  | Mean of all features |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

Dataset to create for final selection

1. Raw dataset with one sound file as one observation – all features & not scaled
2. Raw dataset with one sound file as one observation – all features & scaled
3. Raw dataset with one sound file as one observation & Hilbert transformed - all features & not scaled
4. Raw dataset with one sound file as one observation & Hilbert transformed - all features & scaled
5. Raw dataset with one sound file split into 5 second samples with overlap with each sample as one observation – all features & not scaled
6. Raw dataset with one sound file split into 5 second samples with overlap with each sample as one observation – all features & scaled
7. Raw dataset with one sound file split into 5 second samples with overlap with each sample as one observation & Hilbert transformed - all features & not scaled
8. Raw dataset with one sound file split into 5 second samples with overlap with each sample as one observation & Hilbert transformed - all features & scaled
9. Raw dataset with one sound file split into 1 second samples with overlap with each sample as one observation – all features & not scaled
10. Raw dataset with one sound file split into 1 second samples with overlap with each sample as one observation – all features & scaled
11. Raw dataset with one sound file split into 1 second samples with overlap with each sample as one observation & Hilbert transformed - all features & not scaled
12. Raw dataset with one sound file split into 1 second samples with overlap with each sample as one observation & Hilbert transformed - all features & scaled
13. Mean & IQR dataset with one sound file as one observation – all features & not scaled
14. Mean & IQR dataset with one sound file as one observation – all features & scaled
15. Mean & IQR dataset with one sound file as one observation & Hilbert transformed - all features & not scaled
16. Mean & IQR dataset with one sound file as one observation & Hilbert transformed - all features & scaled
17. Mean & IQR dataset with one sound file split into 5 second samples with overlap with each sample as one observation – all features & not scaled
18. Mean & IQR dataset with one sound file split into 5 second samples with overlap with each sample as one observation – all features & scaled
19. Mean & IQR dataset with one sound file split into 5 second samples with overlap with each sample as one observation & Hilbert transformed - all features & not scaled
20. Mean & IQR dataset with one sound file split into 5 second samples with overlap with each sample as one observation & Hilbert transformed - all features & scaled
21. Mean & IQR dataset with one sound file split into 1 second samples with overlap with each sample as one observation – all features & not scaled
22. Mean & IQR dataset with one sound file split into 1 second samples with overlap with each sample as one observation – all features & scaled
23. Mean & IQR dataset with one sound file split into 1 second samples with overlap with each sample as one observation & Hilbert transformed - all features & not scaled
24. Mean & IQR dataset with one sound file split into 1 second samples with overlap with each sample as one observation & Hilbert transformed - all features & scaled