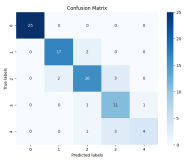
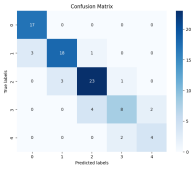
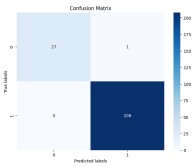
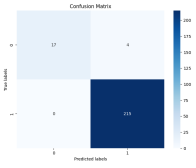
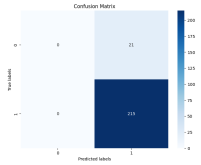
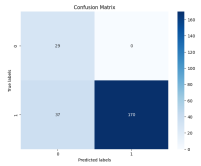
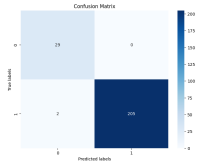
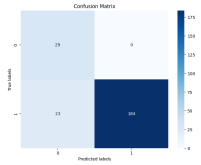
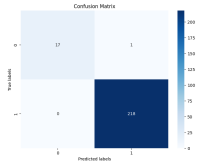
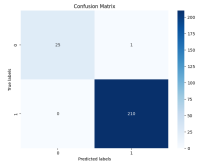
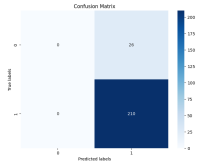
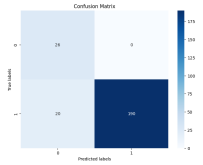
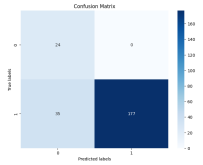
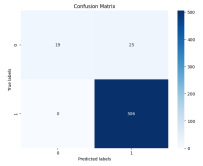
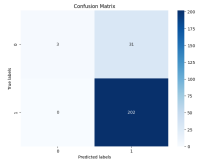
Sl. No.	**Data and features used**	Binary or Multiclass classification	Model	Imbalance in the dataset accounted for?	Confusion Matrix (Vertical axis= True labels, Horizontal axis= Predicted labels)	Comments
1	**PCA of FFT coefficients**	Multiclass	**Random Forest**	N/A	 [25 0 0 0 0 0 17 2 0 0 0 2 16 3 0 0 0 1 11 1 0 0 1 3 4]	**Score: 0.84883721**
2	**PCA of FFT coefficients**	Multiclass	SVC	No	 [17 0 0 0 0 3 18 1 0 0 0 3 23 1 0 0 0 4 8 2 0 0 0 2 4]	Accuracy: 0.81395349
3	**PCA of FFT coefficients after adding 500 random environmental sound samples:** (From ESC 50)	Binary	**Random Forest (50 trees)**	N/A	 [27 1 0 208]	
4	**PCA of FFT coefficients after adding 500 random environmental sound samples:** (From ESC 50)	Binary	**Random Forest (33 trees)**	N/A	 [17 4 0 205]	

5	**PCA of FFT coefficients after adding 500 random environmental sound samples:** (From ESC 50)	Binary	SVC	No	 [0 21 0 215]	Performs very poorly, is unable to identify normal propellers at all
6	**PCA of FFT coefficients after adding 500 random environmental sound samples:** (From ESC 50)	Binary	SVC	Yes	 [29 0 37 170]	Wrongly identified 37 audio samples as normal propeller sound 37 false positives
7	**PCA of FFT coefficients after adding 500 random environmental sound samples:** (From ESC 50)	Binary	KNN	No	 [29 0 2 205]	
8	**PCA of FFT coefficients after adding 500 random environmental sound samples:** (From ESC 50)	Binary	KNN	Yes	 [29 0 23 184]	Wrongly identified 23 audio samples as normal propeller sound 23 false positives
9	**Scaled MFCC coefficients after adding 500 random environmental sound samples:** (From ESC 50)	Binary	**Random Forest (50 trees)**	N/A	 [17 1 0 218]	

10	**Scaled MFCC coefficients after adding 500 random environmental sound samples:** (From ESC 50)	Binary	**Random Forest (33 trees)**	N/A	 [25 1 0 210]	
11	**Scaled MFCC coefficients after adding 500 random environmental sound samples:** (From ESC 50)	Binary	**SVC**	No	 [0 26 0 210]	Performs very poorly, unable to identify normal propellers at all
12	**Scaled MFCC coefficients after adding 500 random environmental sound samples:** (From ESC 50)	Binary	SVC	Yes	 [26 0 20 190]	Wrongly identified 20 audio samples as normal propeller sound
13	**Scaled MFCC coefficients after adding 500 random environmental sound samples:** (From ESC 50)	Binary	KNN	No	[24 0] [0 212]	Accuracy= 1.0
14	**Scaled MFCC coefficients after adding 500 random environmental sound samples:** (From ESC 50)	Binary	KNN	Yes	 [24 0 35 177]	Wrongly predicts 35 audio samples as normal propeller sound
15	**Scaled MFCC coefficients after adding 500 random environmental sound samples:** (From ESC 50)- TRAIN DATASET	Binary	Neural network with Dropout Regularisation	No	 [19 25 0 506]	25 False negatives

16	**Scaled MFCC coefficients after adding 500 random environmental sound samples:** (From ESC 50)- TEST DATASET	Binary	Neural network with Dropout Regularisation	No	 [3 31 0 202]	Performs considerably worse on test data, only able to classify 3 normal propellers correctly, 31 False negatives

Observations:

1. As the dataset itself is imbalanced, (78 normal propeller audio recordings vs 708 recordings of the other class (broken + environmental sounds)), many of the models perform poorly when imbalance is not accounted for using changing of weights or random under/oversampling.
2. RF Regressor is found to be robust to the imbalance in data, and gives better performance as a binary classifier, than when used as a multiclass classifier.
3. The KNN surprisingly gives 100% accuracy on imbalanced data while performing poorly when the imbalance is accounted for.
4. The neural network doesn't achieve very high accuracy on the training data (maybe due to the applied regularisation), but performs considerably poorly on the test dataset, only able to classify 3 normal propellers correctly.
   Clearly, it is affected by the imbalance in the dataset as well as the small size of the dataset itself.