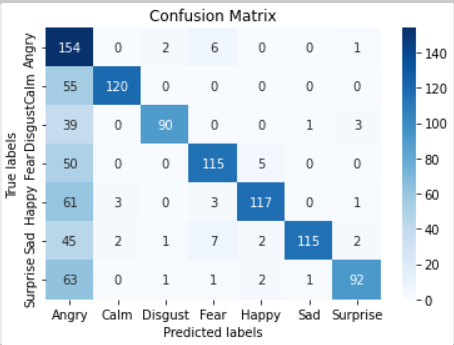
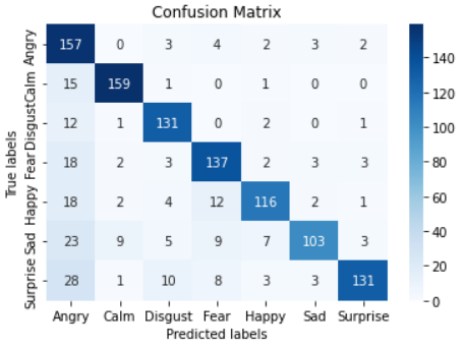
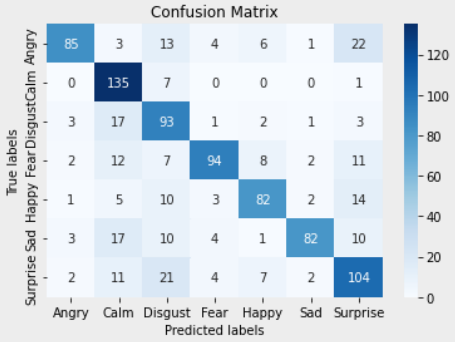
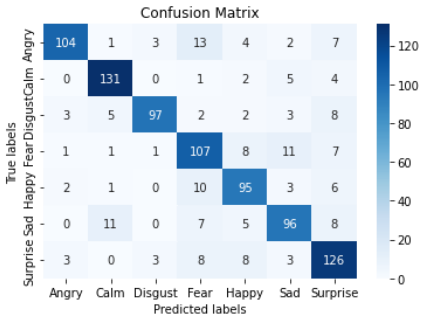
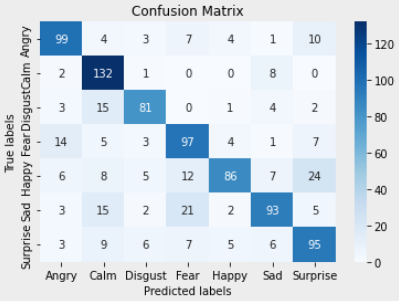
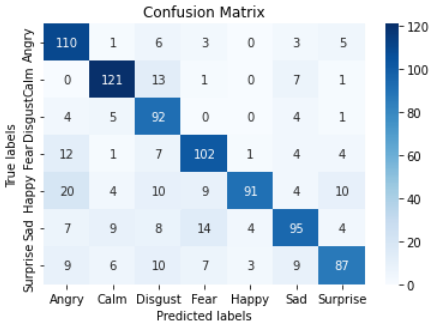
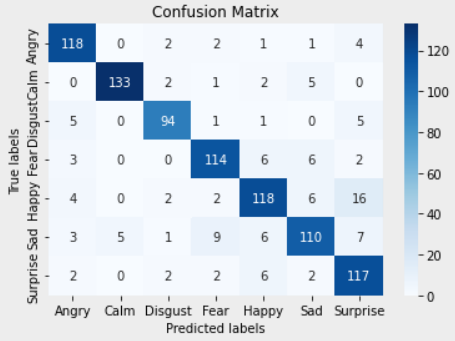
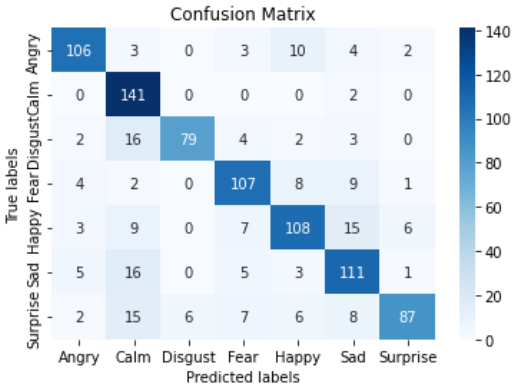
Empathic Art

Model Name	Progress
Baseline CNN	Done
CNN + LSTM	Done
MLP Classifier	Done

Model	Input Feature	Training Time	Epochs	Accuracy	F1-score	Loss
mlpClassifier_1.1	mfcc	26 sec	Alpha = 0.0001	66.12	77.38	4.60
mlpClassifier_1.2	mfcc	45 sec	Alpha = 0.01	76.81	82.95	0.05
baseline_cnn_11	mfcc	9 seconds	11	72.74%	72.83	0.70
baseline_cnn_20	mfcc	2:48	20	81.47%	81.53	0.55
cnn_lstm_11	Mfcc	12 seconds	11	73.60%	73.53	0.66
cnn_lstm_20	Mfcc	22 seconds	20	75.22%	75.02	0.66
cnn_lstm_500_paper	Mfcc	19:36 seconds	500	86.64%	86.78	0.89
cnn_lstm_20_paper	Mfcc	50 seconds	20	79.63%	79.74	0.63

Model	Confusion Matrix
mlpClassifier_1.1	
mlpCLassifier_1.2	

basline_cnn_11	
baseline_cnn_20	
cnn_lstm_11	
cnn_lstm_20	
cnn_lstm_500_paper	

cnn_lstm_20_paper	

Model	Classification Report
mlpClassifier_1.1	

	precision	recall	f1-score	support
Angry	0.86	0.74	0.80	163
Calm	0.96	0.69	0.80	175
Disgust	0.96	0.68	0.79	133
Fear	0.87	0.68	0.76	170
Happy	0.92	0.64	0.75	185
Sad	0.97	0.66	0.79	174
Surprise	0.91	0.60	0.72	160
micro avg	0.92	0.67	0.77	1160
macro avg	0.92	0.67	0.77	1160
weighted avg	0.92	0.67	0.77	1160
samples avg	0.66	0.67	0.67	1160

mlpClassifier_1.2

	precision	recall	f1-score	support
Angry	0.88	0.86	0.87	171
Calm	0.91	0.91	0.91	176
Disgust	0.79	0.92	0.85	147
Fear	0.80	0.85	0.82	168
Happy	0.79	0.80	0.79	155
Sad	0.82	0.68	0.74	159
Surprise	0.88	0.76	0.81	184
micro avg	0.84	0.82	0.83	1160
macro avg	0.84	0.82	0.83	1160
weighted avg	0.84	0.82	0.83	1160
samples avg	0.80	0.82	0.81	1160

baseline_cnn_11

	precision	recall	f1-score	support
Angry	0.89	0.63	0.74	134
Calm	0.68	0.94	0.79	143
Disgust	0.58	0.78	0.66	120
Fear	0.85	0.69	0.76	136
Happy	0.77	0.70	0.74	117
Sad	0.91	0.65	0.76	127
Surprise	0.63	0.69	0.66	151
accuracy			0.73	928
macro avg	0.76	0.73	0.73	928
weighted avg	0.76	0.73	0.73	928

baseline_cnn_20	

	precision	recall	f1-score	support
Angry	0.92	0.78	0.84	134
Calm	0.87	0.92	0.89	143
Disgust	0.93	0.81	0.87	120
Fear	0.72	0.79	0.75	136
Happy	0.77	0.81	0.79	117
Sad	0.78	0.76	0.77	127
Surprise	0.76	0.83	0.79	151
accuracy			0.81	928
macro avg	0.82	0.81	0.82	928
weighted avg	0.82	0.81	0.82	928

cnn_lstm_11	

	precision	recall	f1-score	support
Angry	0.76	0.77	0.77	128
Calm	0.70	0.92	0.80	143
Disgust	0.80	0.76	0.78	106
Fear	0.67	0.74	0.71	131
Happy	0.84	0.58	0.69	148
Sad	0.78	0.66	0.71	141
Surprise	0.66	0.73	0.69	131
accuracy			0.74	928
macro avg	0.75	0.74	0.74	928
weighted avg	0.75	0.74	0.73	928

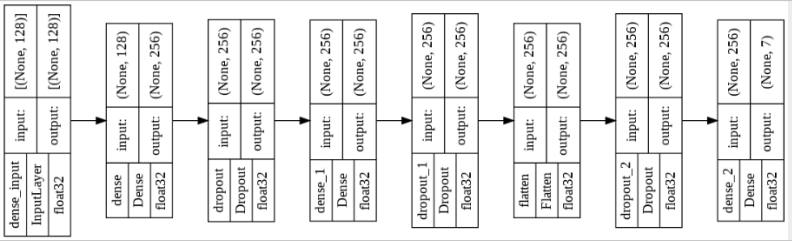
cnn_lstm_20	cnn_lstm_11

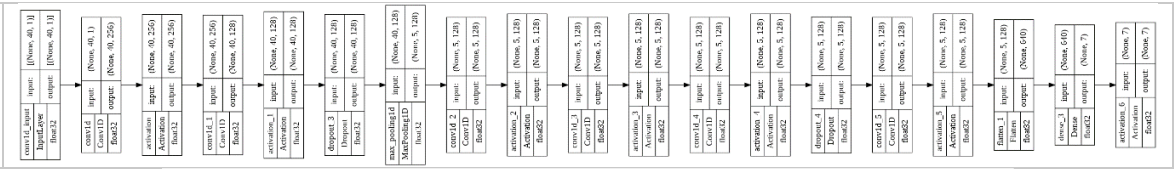
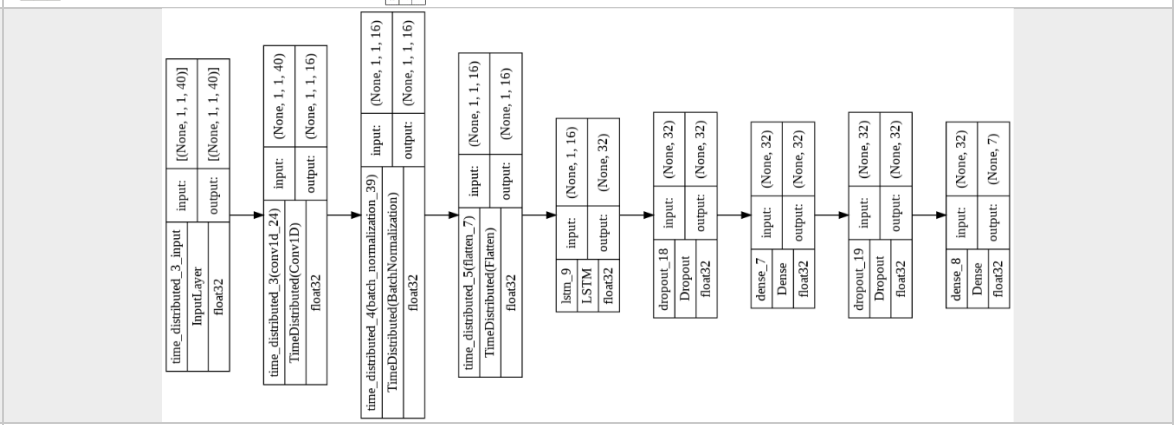
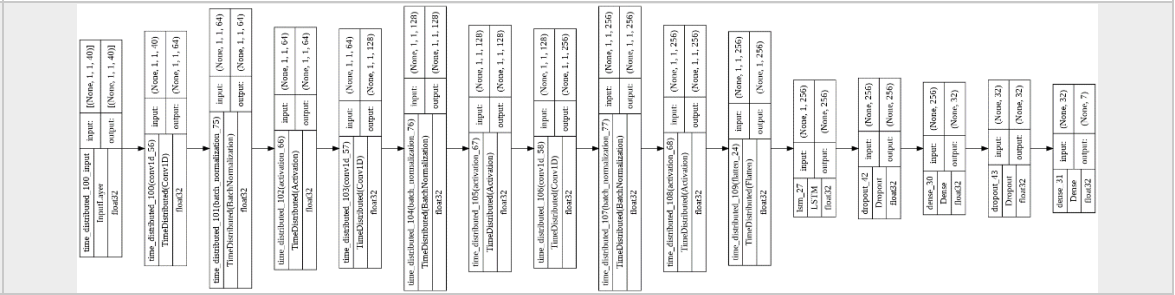
cnn_lstm_500_paper	

	precision	recall	f1-score	support
Angry	0.87	0.92	0.90	128
Calm	0.96	0.93	0.95	143
Disgust	0.91	0.89	0.90	106
Fear	0.87	0.87	0.87	131
Happy	0.84	0.80	0.82	148
Sad	0.85	0.78	0.81	141
Surprise	0.77	0.89	0.83	131
accuracy			0.87	928
macro avg	0.87	0.87	0.87	928
weighted avg	0.87	0.87	0.87	928

cnn_lstm_20_paper	

	precision	recall	f1-score	support
Angry	0.87	0.83	0.85	128
Calm	0.70	0.99	0.82	143
Disgust	0.93	0.75	0.83	106
Fear	0.80	0.82	0.81	131
Happy	0.79	0.73	0.76	148
Sad	0.73	0.79	0.76	141
Surprise	0.90	0.66	0.76	131
accuracy			0.80	928
macro avg	0.82	0.79	0.80	928
weighted avg	0.81	0.80	0.80	928

Model	Architecture
baseline_cnn_11	

baseline_cnn_20	
cnn_lstm_11	
cnn_lstm_20	cnn_lstm_11
cnn_lstm_500_paper	
cnn_lstm_20_paper	cnn_lstm_500_paper

Resources

baseline_cnn_20 architecture: https://github.com/harish24g/Speech-Emotion-Recognition

cnn_lstm_paper_500 architecture: https://pub.towardsai.net/speech-emotion-recognition-ser-using-cnn-and-lstms-4a5dc4c314fd