Assignment 3

Blake Vente

May 2, 2022

Feature Analysis

I chose to perform speaker-wise standardization using z-score normalization, fixing the means at 0 and the standard deviations at 1. The error bars denote the variance within all the samples of a particular emotion

1. An unfounded assumption I had was that I expected for neutral to be close to the mean for all feature values. In fact `neutral` is only close to the mean for `Min Intensity` and that isn't even special because many features had `Min Intensity` close to the mean value.

2. `Min Intensity` caught my eye because as previously mentioned, there is a low standard deviation from the mean: the features' `Min Intensity` values are all closer. This is intuitive because this corresponds with the fact that every sample utterance has a place where a quiet phone is uttered – maybe a fricative – which barely higher than the noise floor. Perhaps the `Min Intensity` actually picked up is actually noise. Regardless, it does not seem like a useful feature to distinguish all features because it's relatively uniform.

3. `hot-anger` has a strikingly low value `Min Intensity` speaker normalized and non-speaker-normalized and a high standard deviation. If anything, I would expect min intensity to be higher than average. This had me stumped for a while, so I listened to many audio samples. I observed that `hot-anger` caused a lot of people to release bursts of air onto the microphone. This actually resulted in really quiet sounds compared to the rest of the audio clips. This is my conjecture for this observation.

4. `despair` was always within 0.2 standard deviations of the mean. While I had expected for `neutral` to be the most "average" emotion, the reality is that `despair` was. Upon further consideration, I found this conclusion satisfying. I view despair as hopelessness characterized by lack of energy (low emotional arousal), causing the speaker's articulations never to drift too far from a baseline.

5. The three leaders in `Min Pitch`, in order are `panic`, `hot-anger`, and `elation`, while the three leaders in `Max Pitch` are very similar `hot-anger`, `panic`, and `elation`. This corresponds with my intuition and is consistent with my discussion of `despair`. These are the emotions which I subjectively remark as having the highest emotional arousal behind them, even if their valance is different. `despair` meant low arousal which meant close to the mean. These three emotions have high emotional arousal which means closer to the extremes.

Classification Experiments

I first looked at my corpus and determined how balanced my data was. Unfortunately, `neutral` had far fewer samples than any other emotion. If I had control over data collection, I would ensure we collected even samples for every emotion, and the same number for every speaker.

Table 1: Frequency Table of corpus labels.

emotion	frequency
contempt	180
happy	177

emotion	frequency
interest	176
despair	174
disgust	172
anxiety	170
elation	159
boredom	154
cold-anger	154
sadness	151
pride	150
shame	148
panic	141
hot-anger	139
neutral	79

We also can observe an over-representation of certain speakers. Ideally this would also be balanced.

Table 2: Frequency table denoting number of samples per speaker.
`df.groupby("speaker").size()`

Speaker	Number of instances
cc	265
cl	368
gg	420
jg	273
mf	299
mk	397
mm	302

I chose `RandomForestClassifier` because *Hybrid Acoustic-Lexical Deep Learning Approach for Deception Detection* by Gideon Mendels, Sarah Ita Levitan, Kai-Zhan Lee, and Julia Hirschberg also used OpenSMILE IS09 features and found it competitive.

Error Analysis

My best classifier reached an aggregate score of 20.9 percent accuracy and 19.8 f_1 score. I found that the best performance was reached by testing on speaker `mm`. I produced the following classification score and confusion matrix.

1. I observe that `despair` utterances perform the worst on average. This is understandable given that earlier, I noted that `depair` had consistently average results. Since random forest picks out combinations of rulse,

Table 3: Table of classification results, averaged across all runs

	precision	recall	f1-score	support
anxiety	0.1774	0.2326	0.1861	24.2857
boredom	0.2465	0.2755	0.2464	22
cold-anger	0.1793	0.0991	0.122	22
contempt	0.2156	0.2153	0.1996	25.7143
despair	0.1193	0.089	0.0937	24.8571
disgust	0.225	0.3229	0.1805	24.5714
elation	0.2058	0.2627	0.2279	22.7143
happy	0.2191	0.3203	0.2537	25.2857
hot-anger	0.3903	0.4728	0.4235	19.8571
interest	0.1675	0.1566	0.1566	25.1429
neutral	0.2571	0.075	0.1079	11.2857
panic	0.3979	0.3039	0.3367	20.1429
pride	0.1468	0.1139	0.1226	21.4286
sadness	0.0867	0.0448	0.0543	21.5714
shame	0.2409	0.2917	0.2543	21.1429
—	—	—	—	—
accuracy	0.209	—	—	—
macro avg	0.2184	0.2184	0.1977	332
weighted avg	0.2267	0.209	0.2004	332

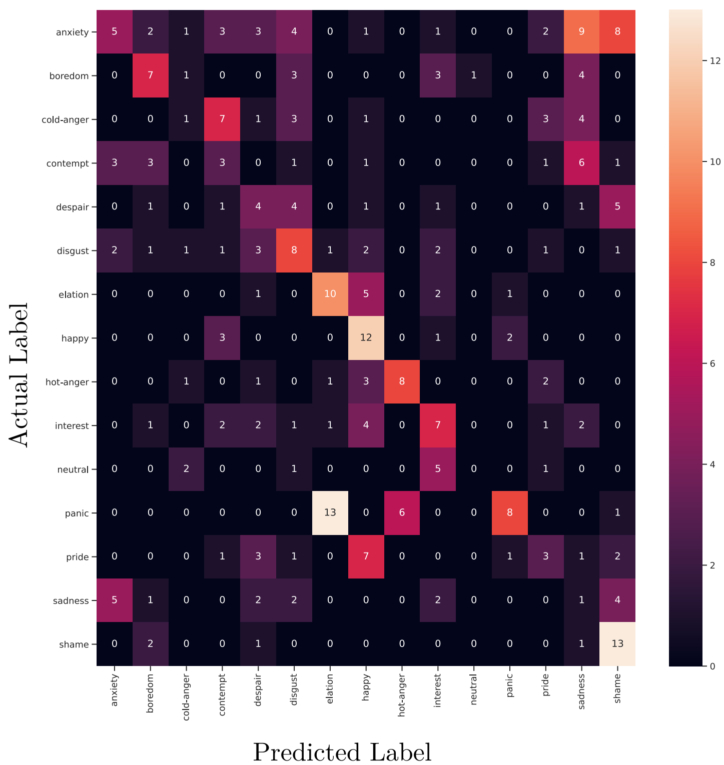
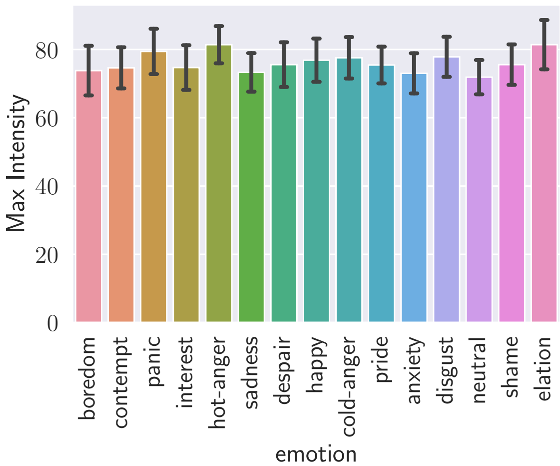
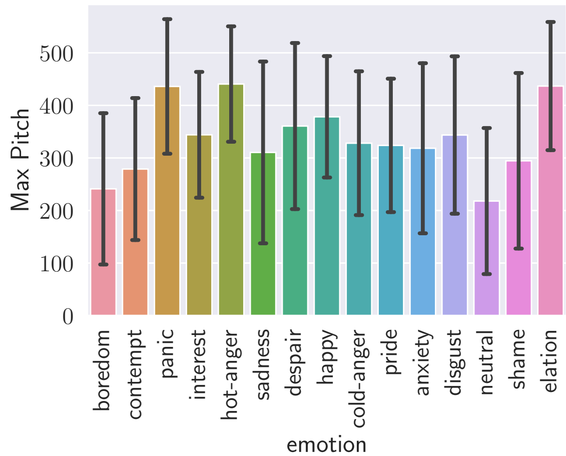
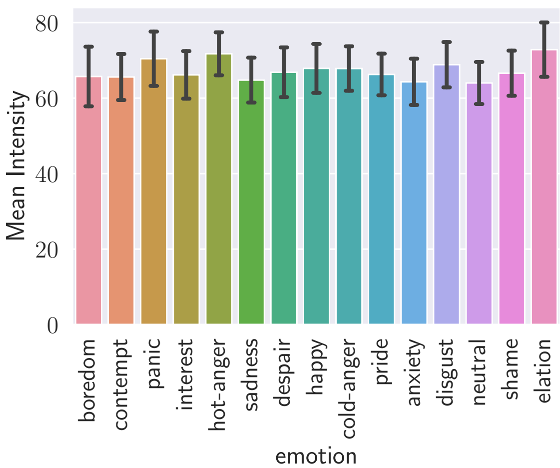


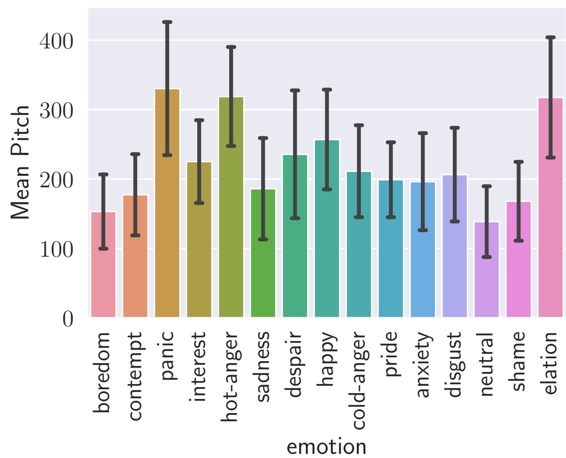
Figure 1: Confusion Matrix of best classifier

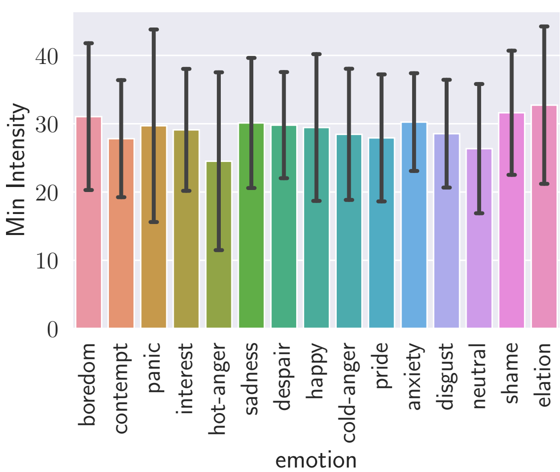
Plots

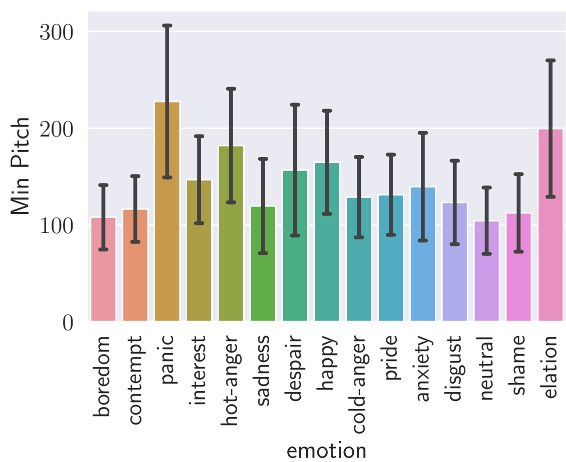

Max Intensity
0
20
40
60
80
boredom
contempt
panic
interest
hot-anger
sadness
despair
happy
cold-anger
pride
anxiety
disgust
neutral
shame
elation
emotion

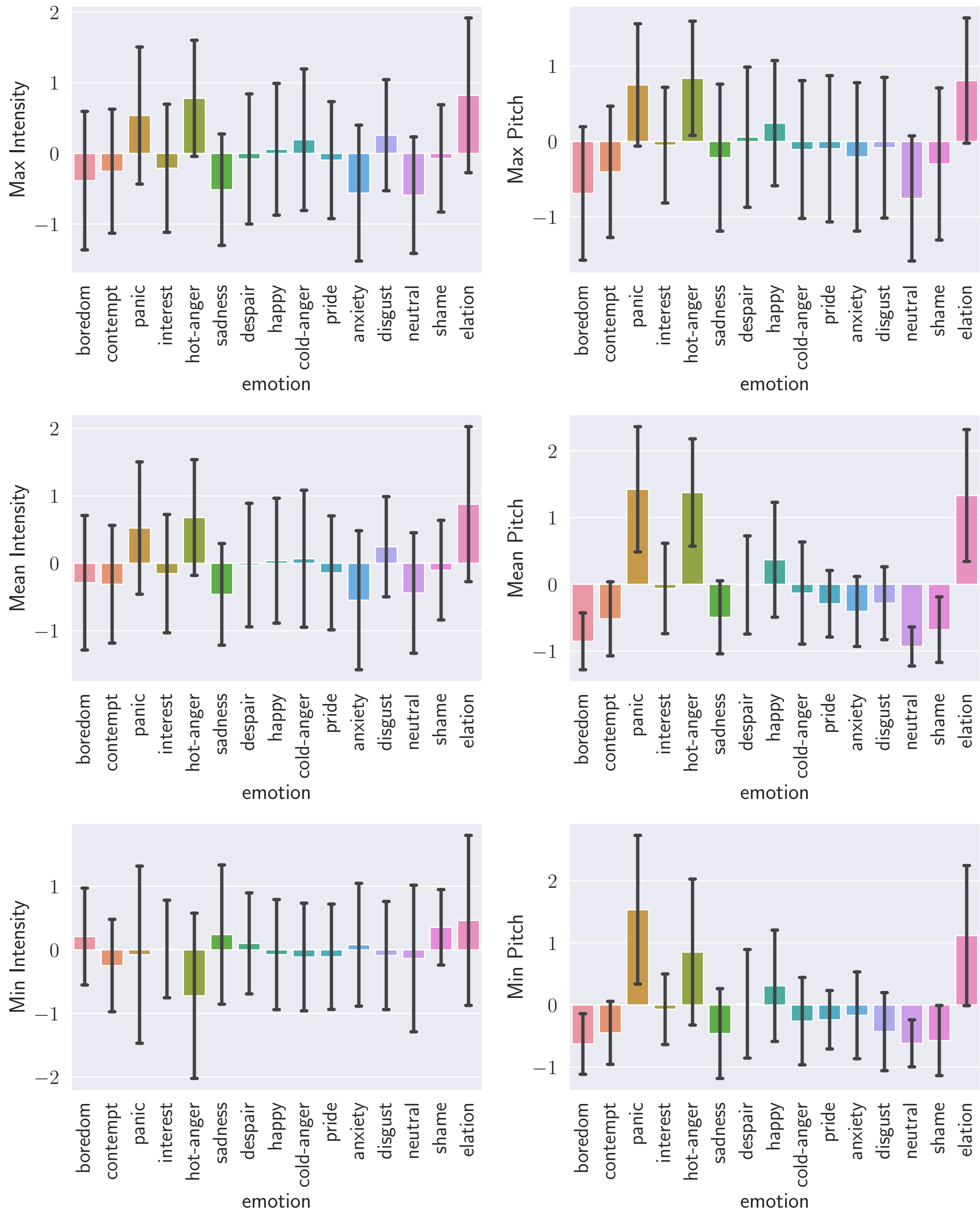

Max Pitch
0
100
200
300
400
500
boredom
contempt
panic
interest
hot-anger
sadness
despair
happy
cold-anger
pride
anxiety
disgust
neutral
shame
elation
emotion


Mean Intensity
0
20
40
60
80
boredom
contempt
panic
interest
hot-anger
sadness
despair
happy
cold-anger
pride
anxiety
disgust
neutral
shame
elation
emotion


Mean Pitch
0
100
200
300
400
boredom
contempt
panic
interest
hot-anger
sadness
despair
happy
cold-anger
pride
anxiety
disgust
neutral
shame
elation
emotion


Min Intensity
0
10
20
30
40
boredom
contempt
panic
interest
hot-anger
sadness
despair
happy
cold-anger
pride
anxiety
disgust
neutral
shame
elation
emotion


Min Pitch
0
100
200
300
boredom
contempt
panic
interest
hot-anger
sadness
despair
happy
cold-anger
pride
anxiety
disgust
neutral
shame
elation
emotion


Max Intensity
Max Pitch
Mean Intensity
Mean Pitch
Min Intensity
Min Pitch
emotion
boredom
contempt
panic
interest
hot-anger
sadness
despair
happy
cold-anger
pride
anxiety
disgust
neutral
shame
elation