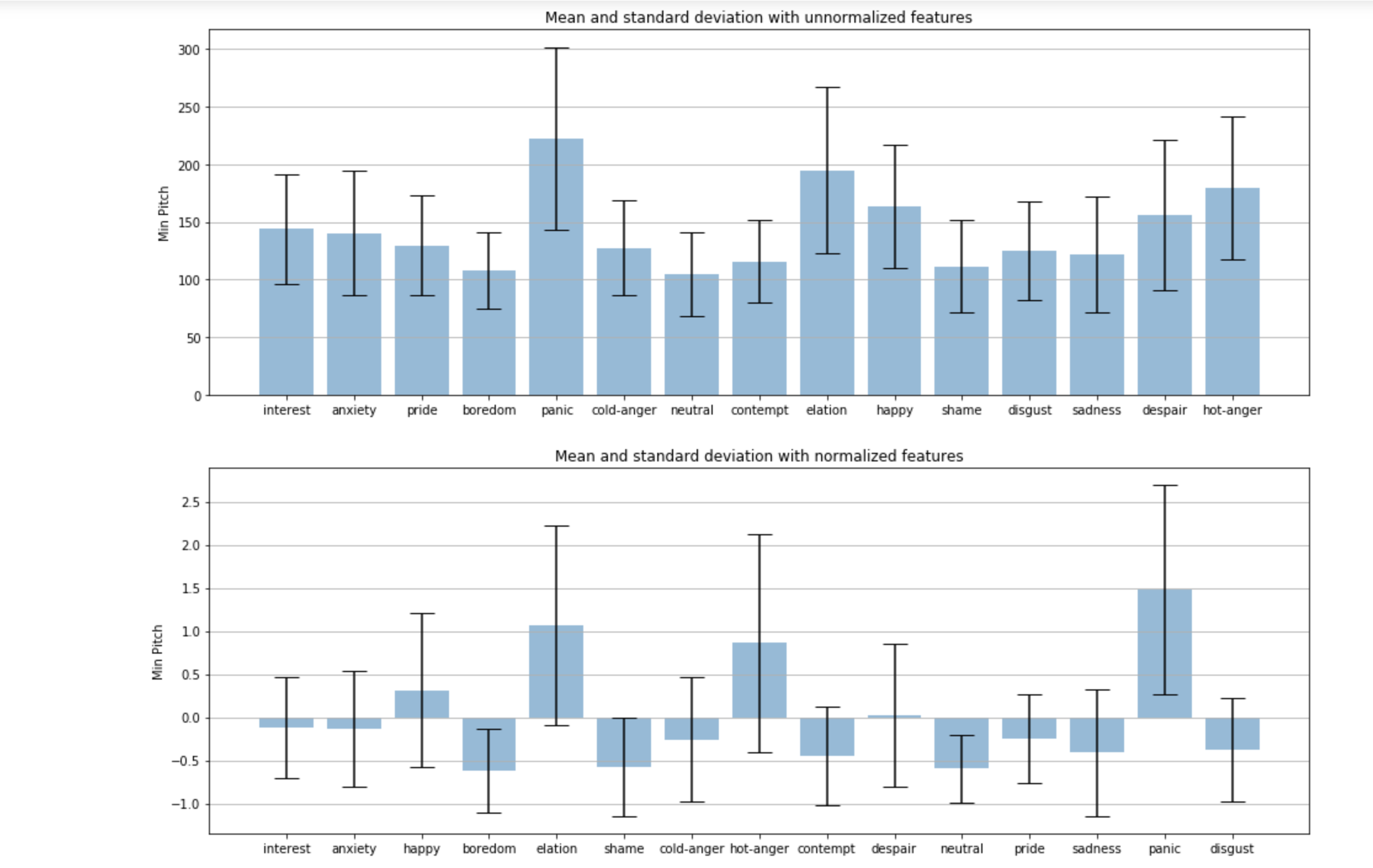
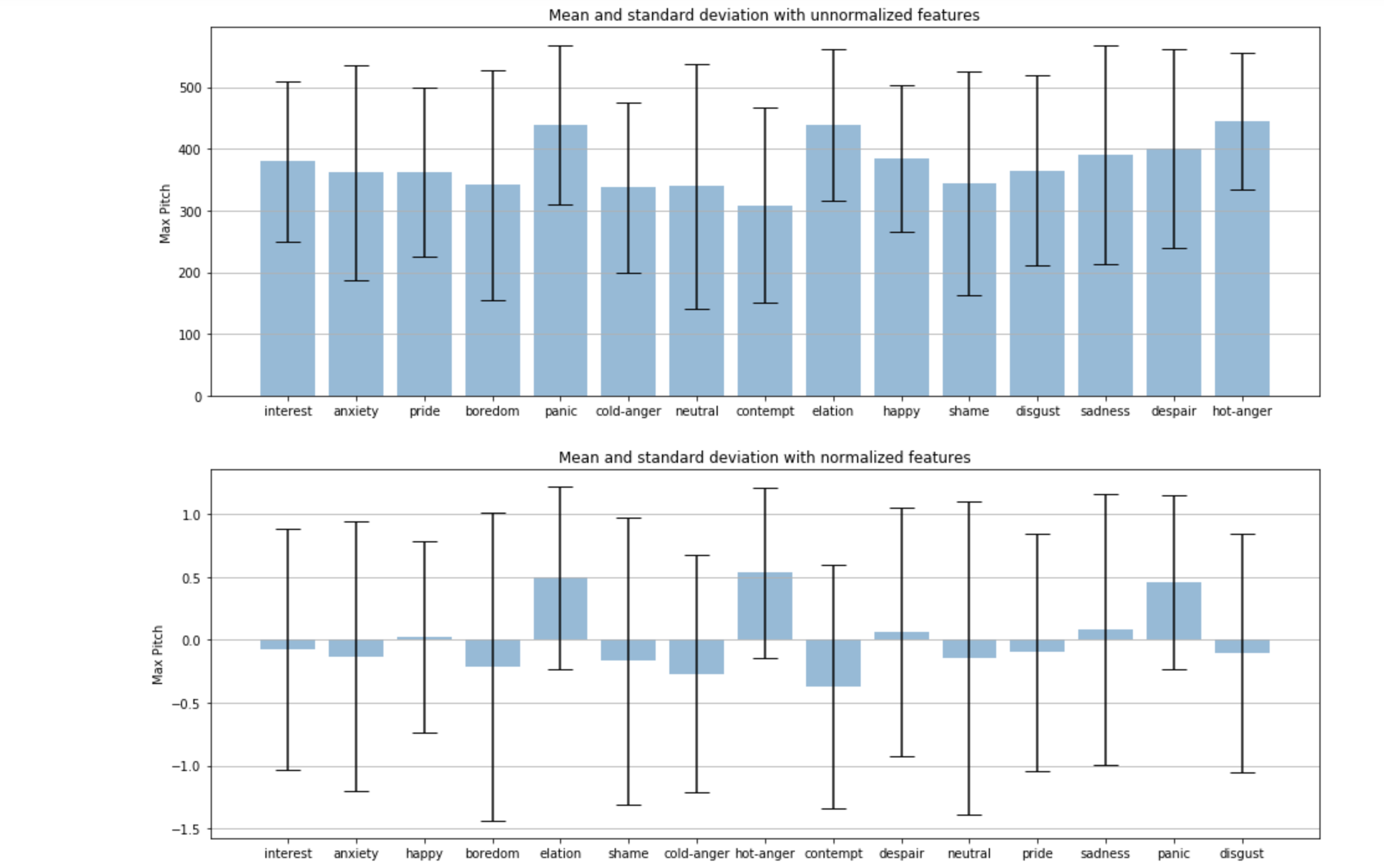
**Task 1: Feature Analysis**

Z- score normalization over each individual speaker was carried out for all the 6 features. The following plots show the visualization of the mean and standard deviation for each feature for both ‘normalized’ and ‘un-normalized’ data.

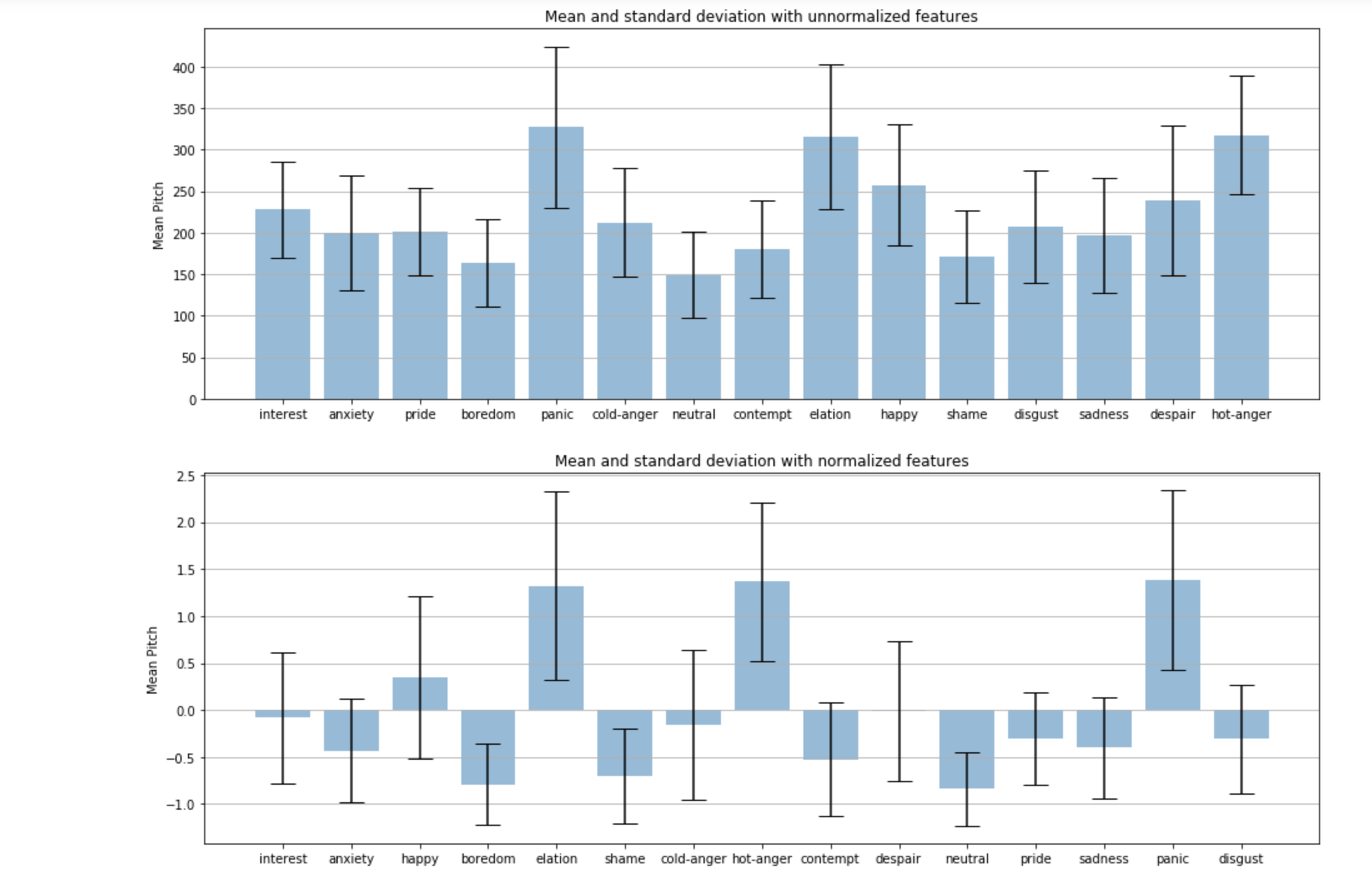
1. **Min Pitch**

****

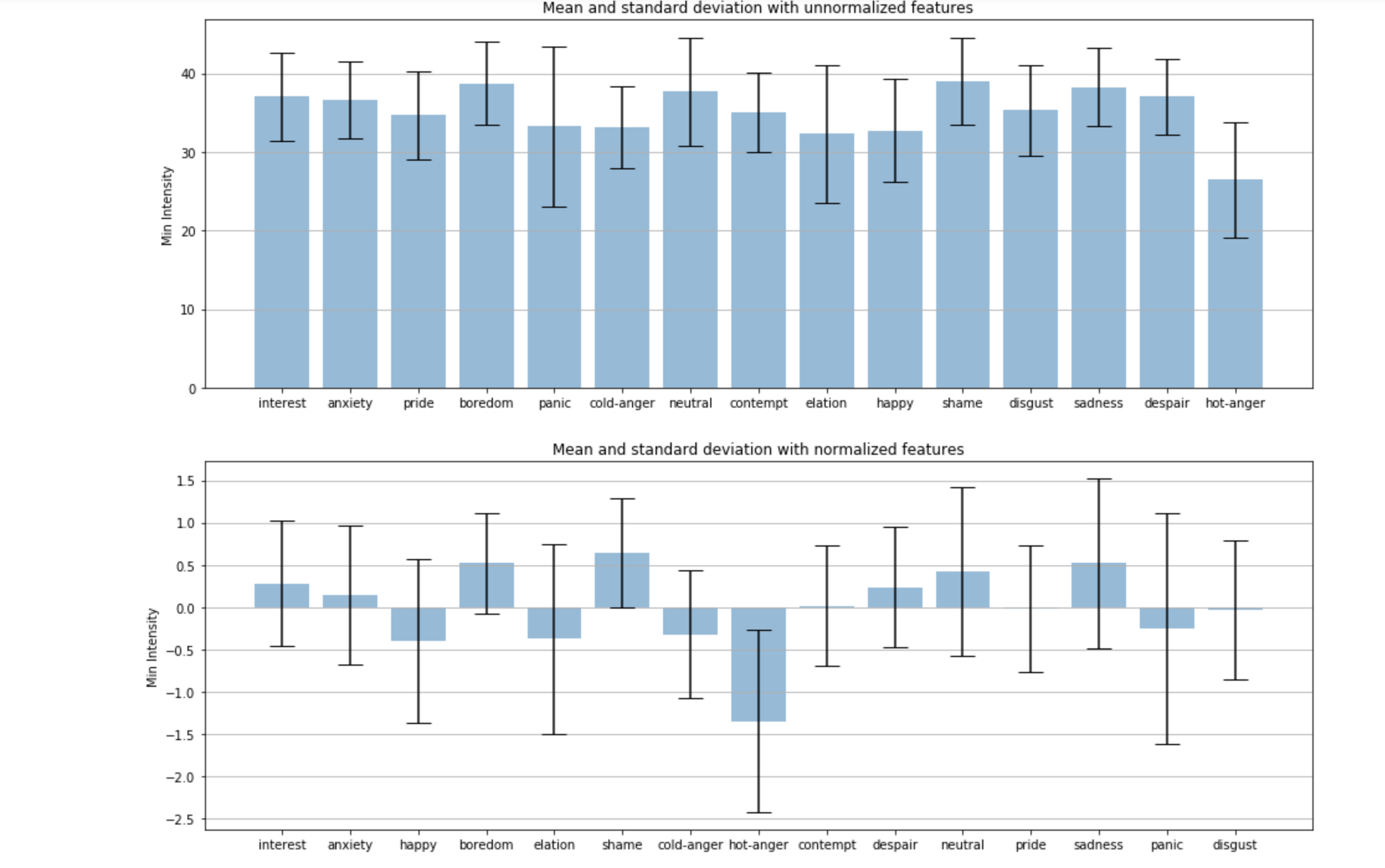
1. **Max Pitch**

****

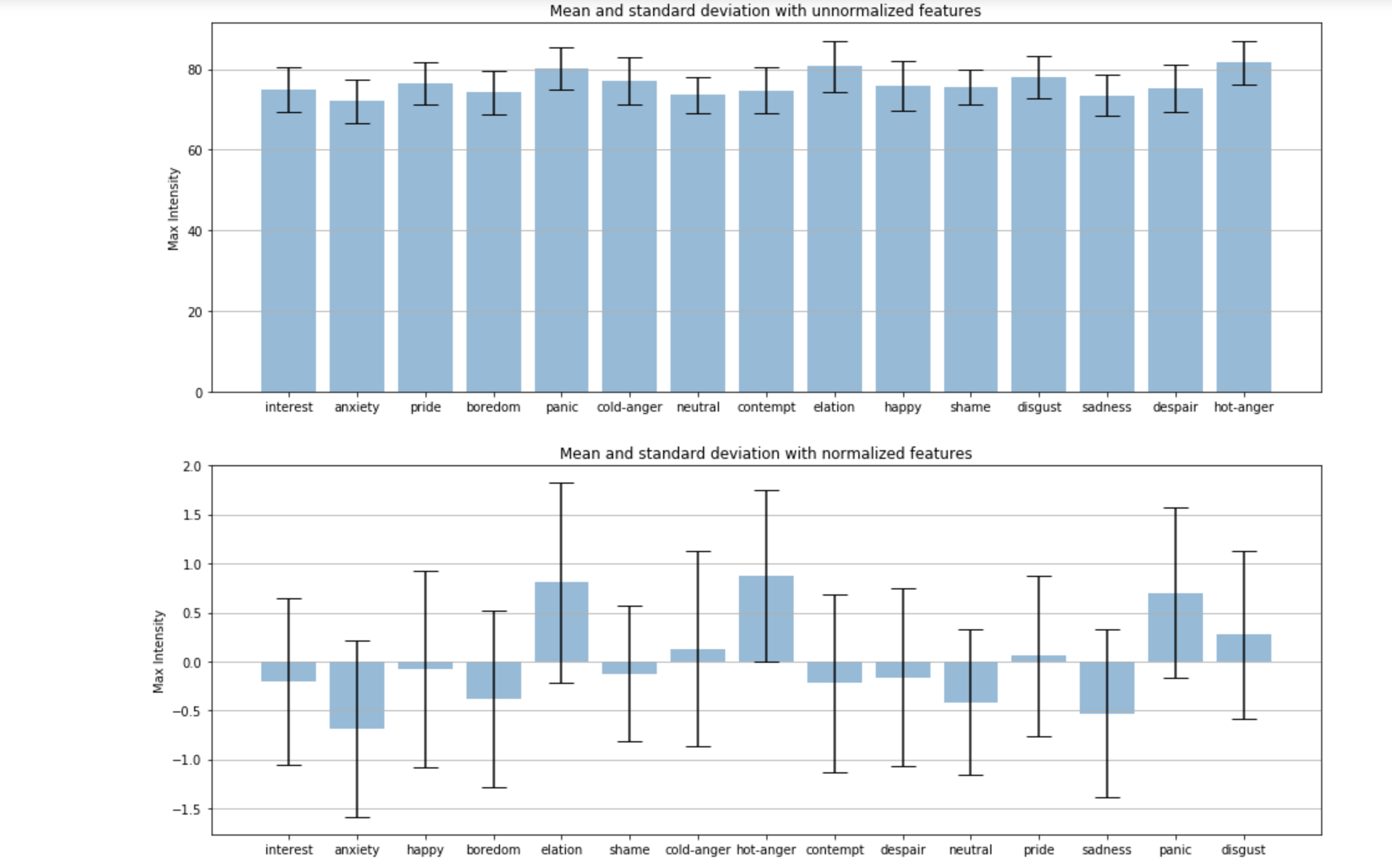
1. **Mean Pitch**

****

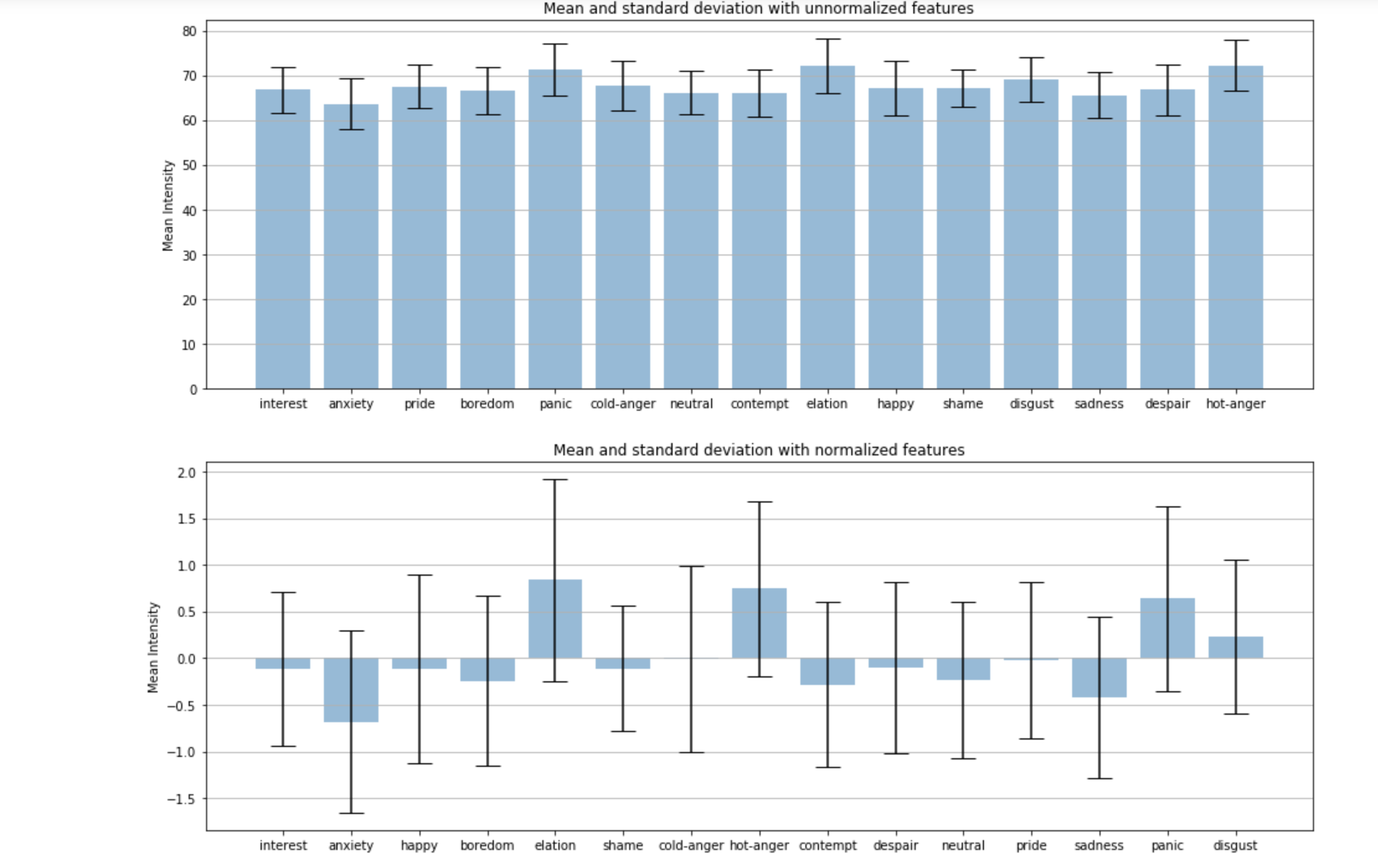
1. **Min Intensity**

****

1. **Max Intensity**

****

1. **Mean Intensity**

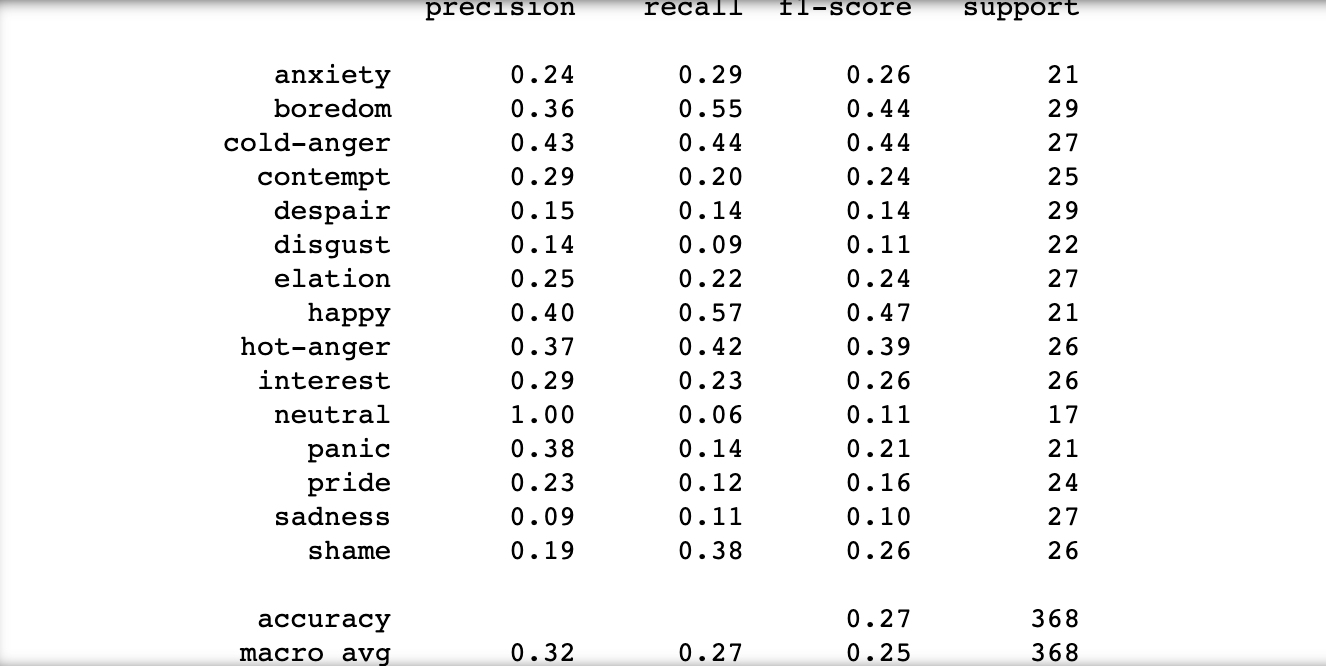
****

**The following are some interesting observations that can be derived from the above plots:**

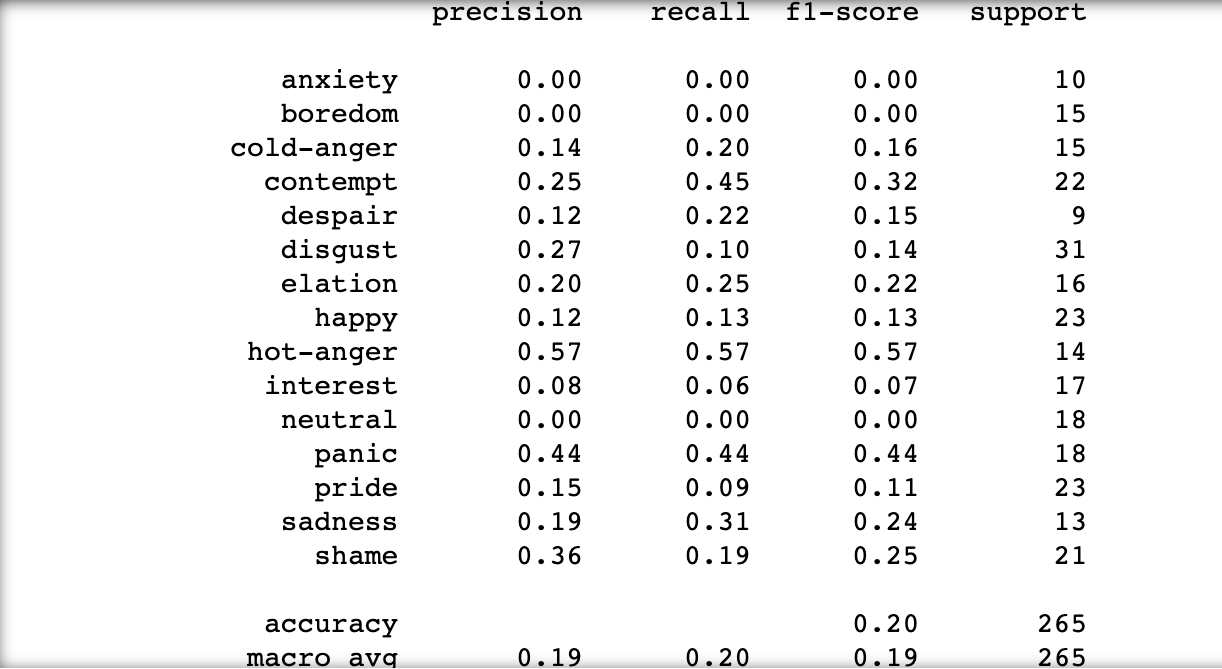
1. Mean pitch is the highest for emotions such as panic, elation, happy and hot-anger as expected from normal intuition. Further, neutral has the lowest minimum pitch.
2. Panic, elation and despair also seem to have the highest standard deviation in pitch values. Whereas hot anger has a very high mean pitch but does not have a very high standard deviation in the pitch values.
3. Mean intensity does not show any significant variation in value among the different emotion categories and hence would not be a very useful feature in differentiating between different emotion categories.
4. Standard deviation for Max Intensity and Mean Intensity is fairly consistent across all features with very little variation.
5. Standard deviation for Max and Mean intensity are much more pronounced and significant in the normalized values as opposed to un-normalized values.

**Task 2: Classification Experiments**

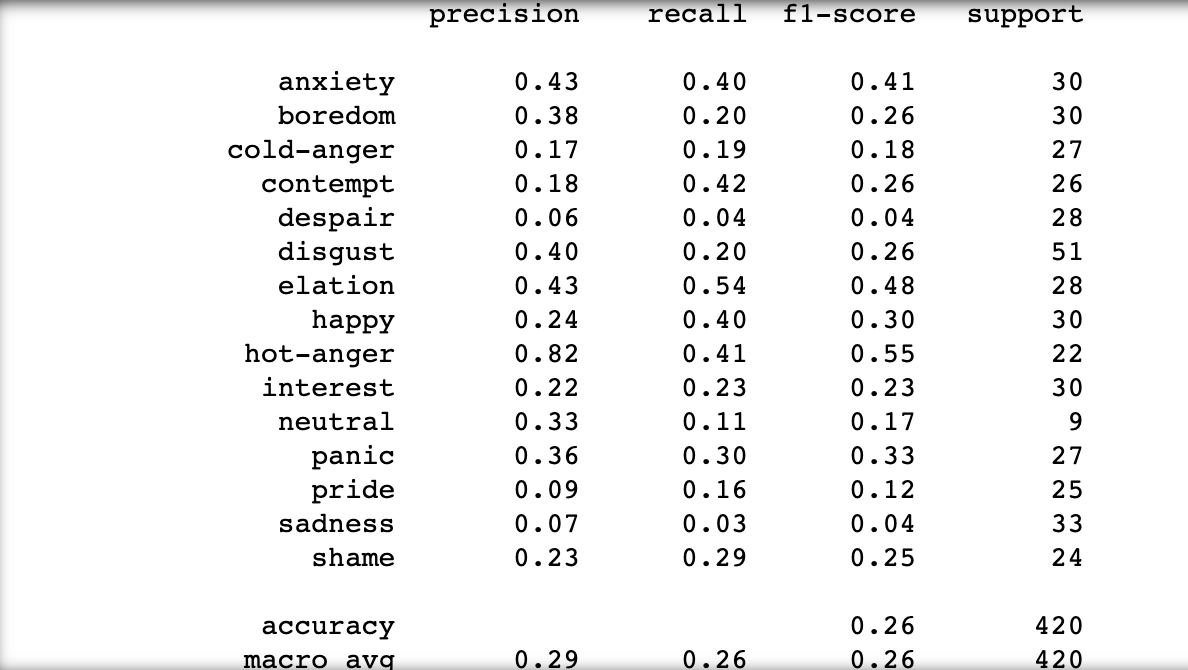
1. **Experiment 1**



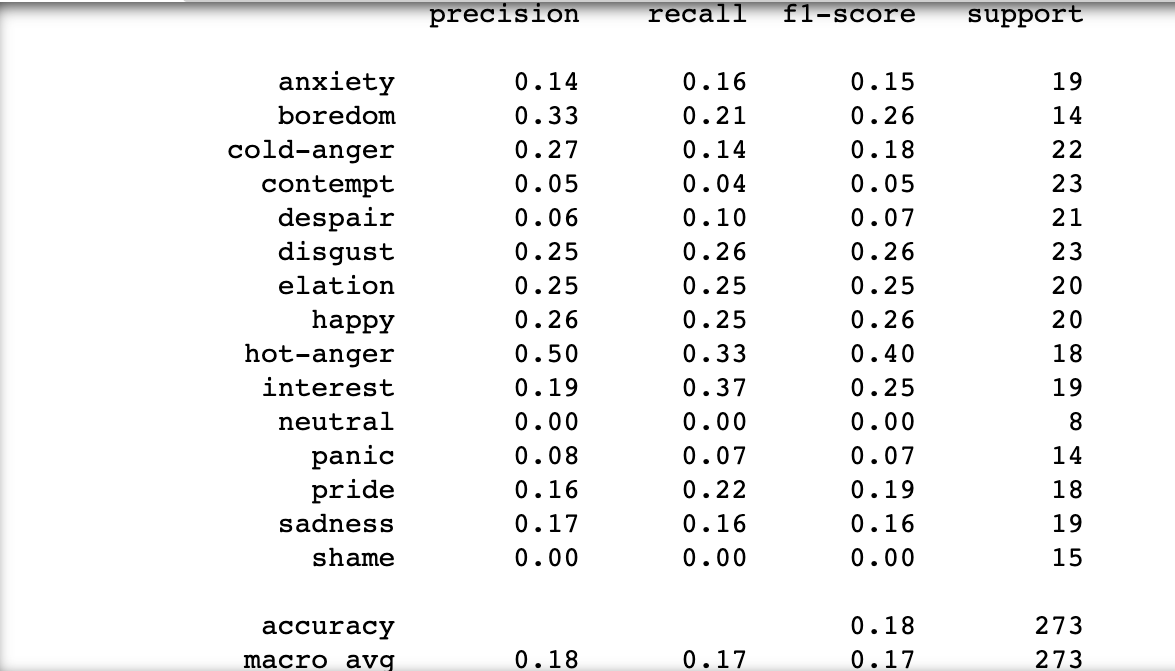
1. **Experiment 2**



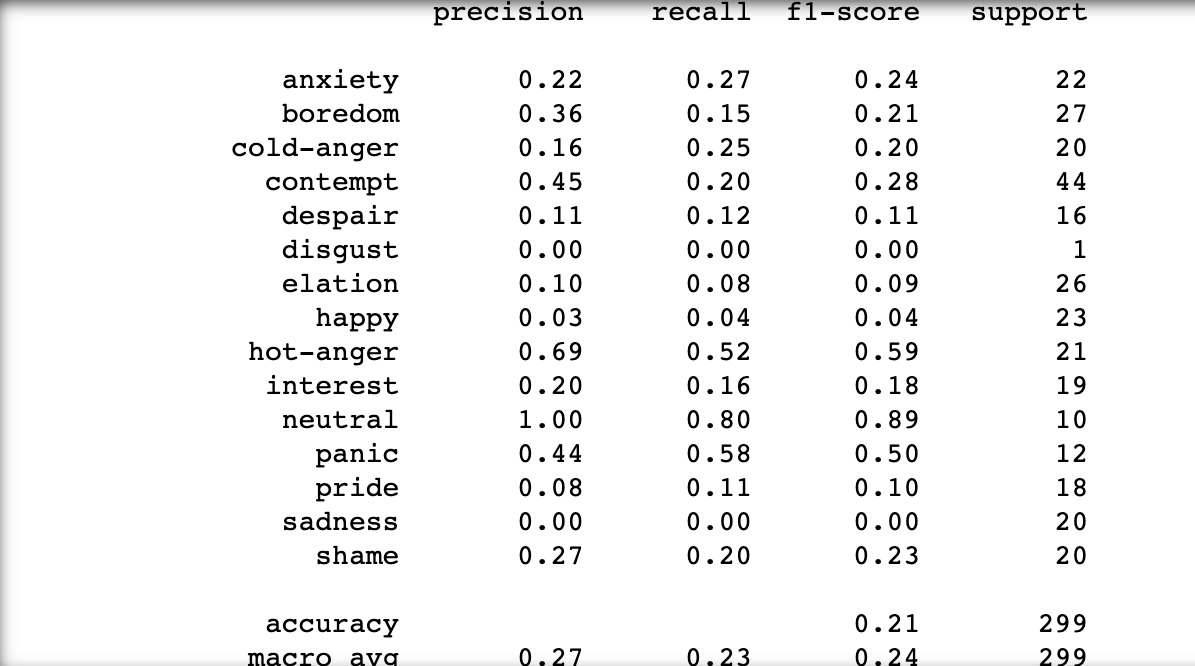
1. **Experiment 3**



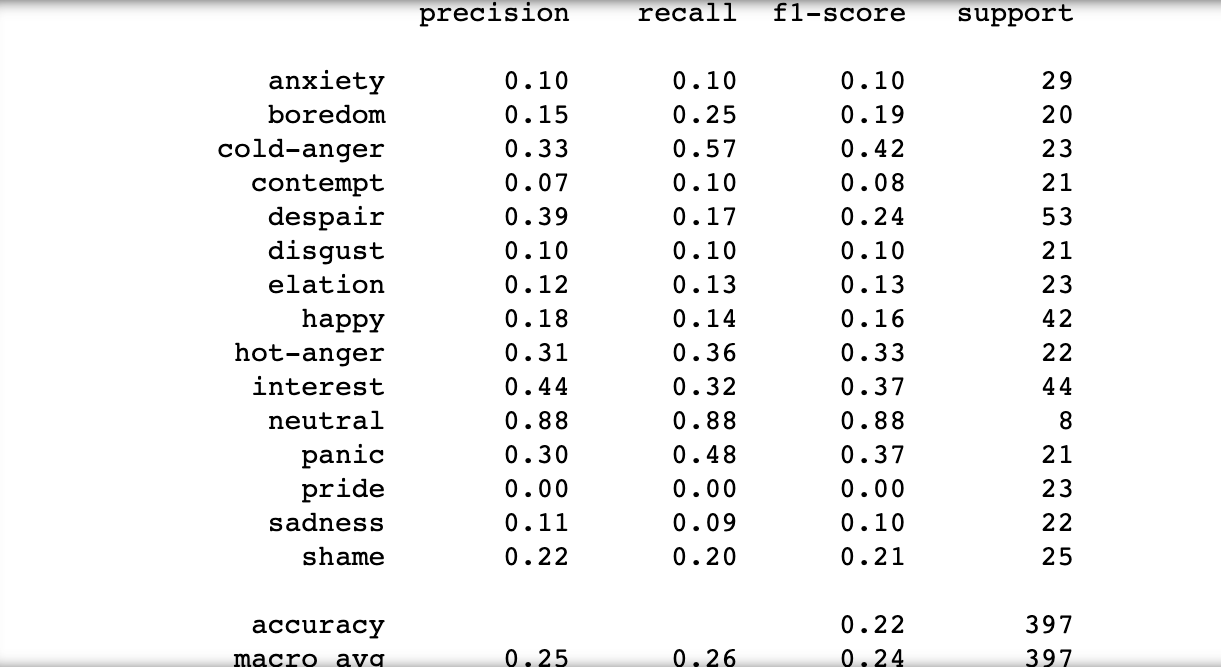
1. **Experiment 4:**



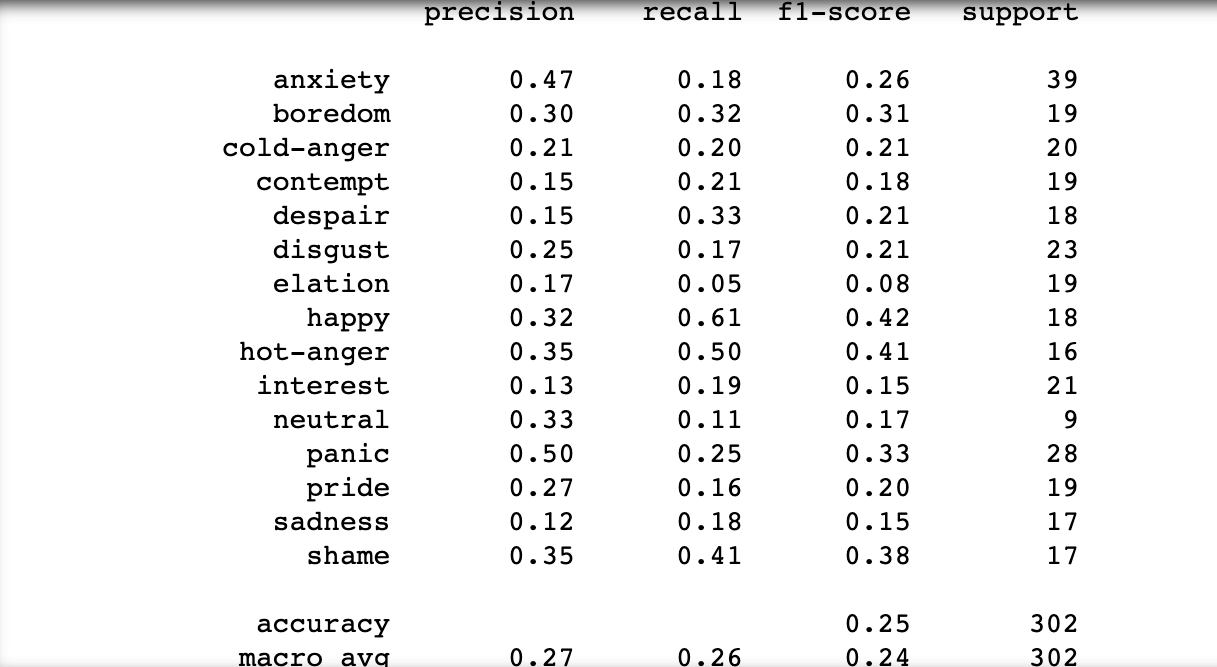
1. **Experiment 5:**



1. **Experiment 6:**



1. **Experiment 7:**



**Task 3**

From the above experiments, it can be seen that the best performing model was for Experiment 1

For Experiment 1, the highest F-1 score was achieved for the emotion category – ‘happy’. The other classes which were easy to predict were ‘boredom’ and ‘cold anger’.

For Experiment 1, the lowest F1- score (i.e. difficult to predict) was for the emotion categories – sadness (F-1 score: 0.10), disgust (F-1 score: 0.11) and neutral (F-1 score: 0.11)

I think category ‘happy’ would have been easier to predict since among the 15 emotion categories it is the only one (other than perhaps pride) which has a positive feeling associated to it and hence on the basis of the acoustic and prosodic features provided, the classifier is able to correctly differentiate the ‘happy’ emotion from other emotions.

Similarly, emotions such as ‘cold anger’, ‘hot anger’ and ‘boredom’ can be relatively easily identified from an individual’s speech with certain distinctive features to be able to spot such emotions. Also, like ‘happy’ they are an extreme emotion category which are in general more easily identifiable in speech.

Emotions such as ‘sadness’, ‘despair’ and ‘disgust’ can be quite similar in terms of being reflected in speech and hence it is highly likely that the classifier confuses between identifying these classes leading to a low f-1 score for their classification.

Also an emotion such as ‘neutral’ would not have any peculiar variations in features such as ‘pitch’ and ‘frequency’, ‘RMS Energy’, making it difficult to identify that emotion.

**Ways to improve:**

1. Feature selection can be performed to take a more useful subset from the set of features currently used to achieve better F-1 scores.
2. Currently, an XGBoost model was used for the classification task. Other models such as deep learning based model can also be tried out with hyper parameter tuning to achieve better results.
3. Further, class balancing can be done to ensure equal representation of the emotion classes in the training data set for each experiment.