

HomeWork 1 Report : Annotation Aggregation and Exploratory Analysis

NAME: Kavya Gengedulla Yeranarasappa

USC ID : 3206274199

1) Fleiss' Kappa Inter-Annotator Agreement

The Fleiss' kappa coefficient was computed using the provided tabulatedVotes.csv dataset to evaluate the level of agreement among multiple raters in categorizing emotions across six predefined categories (A, D, F, H, N, S). Following the guidelines provided in the assignment, the interpretation of the kappa value was applied. The calculated **Fleiss' kappa value is 0.4116** which indicates an inter-rater reliability characterized by **MODERATE AGREEMENT**. This conclusion aligns with the interpretation criteria specified, where a kappa value falling within the range of **0.41 to 0.60** denotes moderate agreement among raters.

```
selected_columns = ['A', 'D', 'F', 'S', 'H', 'N', 'S']
filtered_data = data[selected_columns].to_numpy()
kappa = fleiss_kappa(filtered_data)
print("Fleiss' Kappa:", kappa)
print("Fleiss Kappa value indicates that there is moderate inter annotation agreement")
```

Fleiss' Kappa: 0.41163477995629666

Fleiss Kappa value indicates that there is moderate inter annotation agreement

2) Identification of Audio and Visual Cues

TYPE	AUDIO	VISUAL
ANGRY	Increased volume and intensity in speech	Raised eyebrows
HAPPY	Lively and expression of joy in the tone	Bright and expressive eyes
SAD	Subdued and lack of energy in the talk	Slouched posture

3) Confirmation of Suspicions and Subjective Observations

To confirm the suspicions and subjective observations, 20 videos for each emotion (HAPPY and SAD) were manually annotated for the presence or absence of the identified cues. The annotations were saved in the **emotion_annotation.csv** file and the ground truth values were saved in the **real_annotation.csv** file with three columns: filename, HAPPY (Y or N), and SAD (Y or N). A contingency table for the Chi-Squared test and P-Value was then calculated to analyze the significance of the presence of cues in the videos.

Below is the **contingency table** to analyze the association between the manually annotated HAPPY flags and the real HAPPY flags.

	HAPPY_y		N	Y
HAPPY_x				
	N	17	3	
	Y	3	17	

The Chi-squared statistic and the p-value calculated are as below:

- **Chi-squared statistic: 16.9**
- **P-value: 3.94**

Based on the P-value which is > 0.05 threshold , the null hypothesis is rejected.

This indicates that the presence of cues identified for HAPPY videos is significantly associated with the real HAPPY flags