Advanced EDA of Video Text Dataset

Data Science Team

September 19, 2024

Dashboard

For seeing all the code live interactively,

Visit Dashboard

https://eda-analysis-iby-0.streamlit.app/

Contents

1. Introduction

Introduction

The goal of this report is to find hidden patterns and insights of students based on their transcript, speech patterns and emotional expressions during video presentations.

The dataset contains information about the emotions detected in each frame of the video, the gaze behavior of the students, and the transcript of their speech. By combining these different data sources, we aim to gain insights into the students' communication skills and identify areas of expertise or improvement.

The actionable insights from this analysis include:

- **Recruitment Decision:** Based on the analysis of communication skills, I determined whether a candidate is suitable for a role and ranked them .
  By considering factors such as speech speed, positivity, and confidence,They can be assessed whether one is fit for the role or not.
- **Areas of Expertise:** By analyzing the transcript data, i first identified the students' areas of expertise based on the topics they discuss and the language they use. This information can help in assigning roles or projects that align with their strengths.
- **Decision-Making Insights:**I provided various insights that can help in making decisions about the candidate.
  For example, identifying patterns in emotional expressions or gaze behavior that correlate with effective communication can provide useful guidance for future interactions.

2. Data Overview

Emotion Data Overview

emotion_data: This dataset contains the following columns. We can inspect the first few rows using *emotion_df.head()*.

Emotion Data Sample

This is a sample of emotion data. **Description:** This dataset represents the emotions detected at each timestamp of the video, along with the dominant emotion for each image sequence.

Image Seq	Angry	Disgust	Fear	Happy	Sad	Surprise	Neutral	Dominant Emotion
0	4.32	0.00	2.88	1.65	2.78	0.60	87.77	Neutral
1	53.23	2.98	12.74	1.52	1.05	27.22	1.26	Angry
2	8.80	0.03	2.97	16.83	39.88	0.28	31.21	Sad
3	9.45	0.11	1.55	20.93	3.50	0.91	63.54	Neutral
4	56.00	0.00	0.16	5.58	0.20	12.81	25.25	Angry

Structure of Gaze Data

Gaze Data Structure

Image Seq	Gaze	Blink	Eye Offset
1	1	0	6.23
2	1	0	22.73
3	1	0	2.57
4	1	0	21.11
5	1	0	1.85

Transcript Data Sample

Transcript Data Sample

id	text	tokens	positive	negative	neutral	confident	hesitant	concise	enthusiastic	speech_speed
0	Hello, I am J	[50364, 24]	0.5803	0.1523	0.2675	0.8467	0.8457	0.6358	0.6478	2.518
1	IIM Coikode.	[50642, 28]	0.5503	0.1893	0.2604	0.6793	0.7337	0.5441	0.4174	3.2178
2	Technology	[50844, 15]	0.6399	0.1111	0.2490	0.9027	0.8346	0.7159	0.7001	2.8689
3	of three yea	[51088, 29]	0.4419	0.3992	0.1589	0.7743	0.8130	0.5225	0.2799	3.750
4	as a medical	[51288, 38]	0.2363	0.5320	0.2317	0.2860	0.5614	0.3344	0.1973	3.5417

Basic Statistics

Basic Statistics

	id	seek	start	end	positive	negative	neutral	confident	hesitant	concise	enthusiastic	speech_speed
count	18	18	18	18	18	18	18	18	18	18	18	18
mean	8.5	3,009.33	41.0022	45.9311	0.7092	0.1412	0.1496	0.7338	0.4852	0.4294	0.4665	3.1138
std	5.3385	2,598.47	26.117	26.2949	0.2073	0.1549	0.0810	0.2083	0.2608	0.2726	0.2863	0.600
min	0	0	0	5.56	0.2363	0.0050	0.0146	0.2860	0.0084	0.0128	0.0886	2.0349
25%	4.25	0	19.68	24.4	0.5879	0.0433	0.0829	0.5769	0.3429	0.2808	0.2114	2.6057
50%	8.5	2,776	40.56	46.64	0.7397	0.0804	0.1557	0.7899	0.4078	0.4415	0.4189	3.1342
75%	12.75	5,336	62.42	66.66	0.8701	0.1602	0.2246	0.8986	0.7108	0.6129	0.6870	3.5897
max	17	8,272	82.72	88.72	0.9804	0.5320	0.2675	0.9809	0.8457	0.9197	0.9903	4.1667

Note: In the transcript data, the columns *text*, *tokens*, *temperature*, *avg_logprob*, *compression_ratio*, and *no_speech_prob* are removed for simplification.

3. Data Preparation and Integration

So, the data is spread across three different CSV files, and what I need to do is pull out key features from all of them to create a single, clean DataFrame. First off, I'm going to take the **dominant emotions** for each student across the entire video. To do this, I'll average the dominant emotions over every frame and then store them as **dominant_emotion_top_1** and **dominant_emotion_top_2**. These two values will give us a quick idea of the main emotional tones for each student.

1. First, I'll **extract and compute the average values** from the **transcript data**. This is important because we need to analyze how students behave by looking at things like **positivity**, **negativity**, and **speech speed**. By averaging these features across their complete transcript, I can see overall patterns, like how often they seem confident or hesitant.

2. Then, I'll **store these averages in a list of DataFrames**. The reason for this is to keep track of each student's data separately at first. This way, when I combine them later, it's easier to manage and analyze.

3. After that, I'll **combine all the averages into one final DataFrame**. This step is key because it gives me a **well-organized dataset** where I can quickly compare each student's average speech features. So, up to here, the **transcript part** is done.

4. Next, I'll **extract the dominant emotions** from the **emotion data**. The goal here is to capture the emotional tone of each student by identifying the **most frequent emotions**—whether they're happy, sad, or something else. This adds more depth to the average speech features we just looked at.

5. Once I have the dominant emotions, I'll **store the top two** for each student. This is important because it gives us a quick snapshot of their **emotional profile** during speech, covering both the primary and secondary emotions. It's like getting the full emotional picture.

6. I'll then **create a DataFrame** to hold all these dominant emotions. Having this in a separate DataFrame makes it easy to **merge with the average speech features**, so I can analyze everything together in one place.

7. Now, I'll **merge the average features DataFrame with the dominant emotions DataFrame**. This combination gives me a comprehensive dataset, mixing both the **quantitative data** (like positivity and speech speed) with more

qualitative data (like emotions), which lets me dig deeper into how students' behavior and emotions interact.

8. Finally, I'll **display the complete DataFrame**. This final view will give a **clear overview** of each student's **speech patterns** and **emotional tendencies**, which will help understand their **communication behavior** better and could guide future feedback or interventions.

So this is my new DATAFRAME , Let's call it FINAL_DF

Final Dataframe

id	avg_positive	avg_negative	avg_neutral	avg_confident
1	0.709199	0.141214	0.149586	0.733828
2	0.722006	0.107541	0.170453	0.684879

avg_hesitant	avg_concise	avg_enthusiastic	avg_speech_speed
0.485172	0.429418	0.466497	3.113771
0.436158	0.484221	0.516685	3.269092

dominant_emotion_top1	dominant_emotion_top2	gaze_score
neutral	fear	0.625000
happy	neutral	0.609195

blink_sum	eye_offset_mean	eye_offset_max	eye_offset_min
0.000000	15.801362	65.0276	-33.4655
4.597701	21.768546	67.6710	-15.2405

eye_offset_std	image_seq_count
17.858517	88
15.619435	87

4. Data Preprocessing

Steps for Data preprocessing

The following steps were performed to preprocess the data before conducting any Exploratory Data Analysis (EDA):

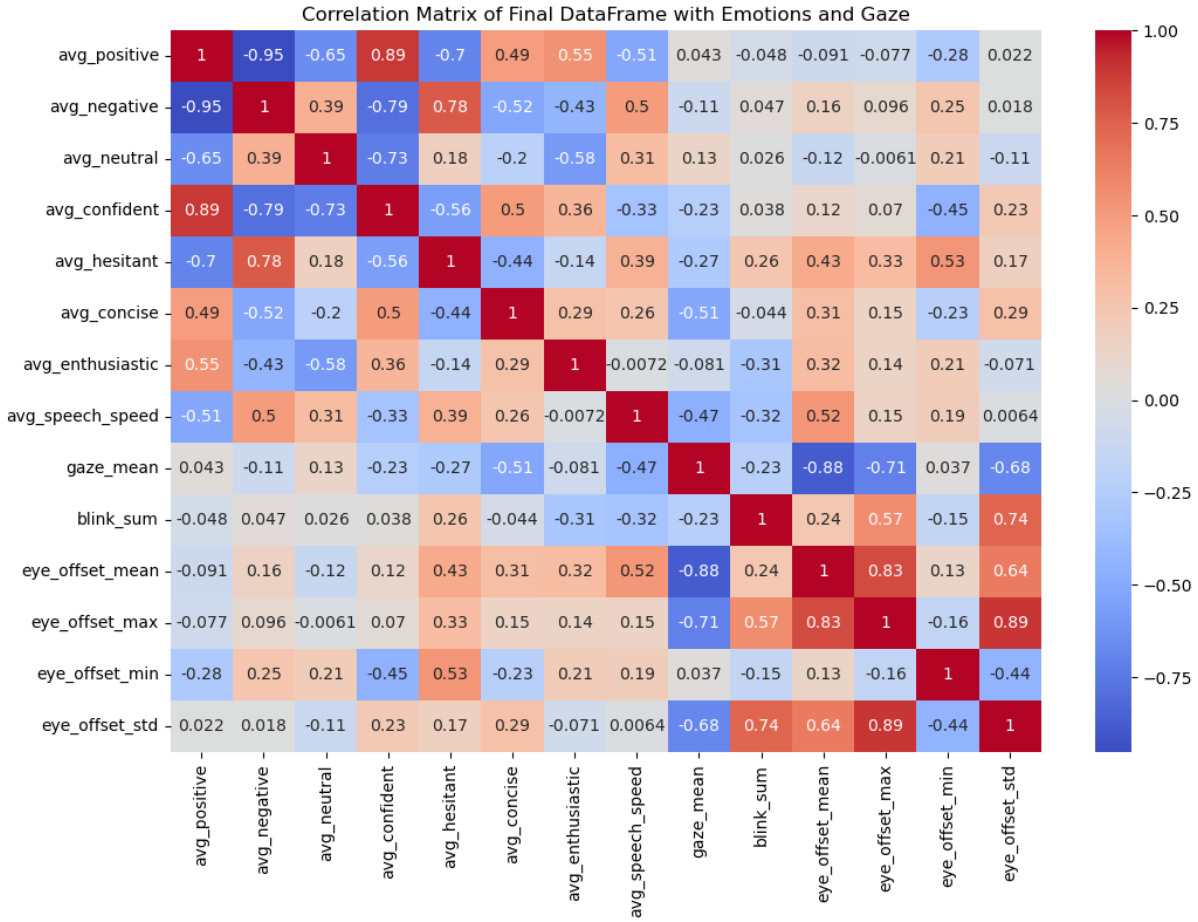
- **Handling Missing Values:** Replaced missing or null values using techniques such as mean, median, or mode imputation based on the feature type.
- **Encoding Categorical Features:** Converted categorical variables into numeric form using label encoding or one-hot encoding.
- **Scaling and Normalization:** Applied standard scaling or min-max normalization to ensure that features with different units or ranges do not dominate the model's learning process.

	avg_positive	avg_negative	avg_neutral	avg_confident	avg_hesitant	avg_concise	avg_enthusiastic	avg_speech_speed
count	10.000000	10.000000	10.000000	10.000000	10.000000	10.000000	10.000000	10.000000
mean	0.652534	0.178566	0.168900	0.635428	0.507159	0.401131	0.446340	2.971076
std	0.058600	0.048346	0.019202	0.060950	0.052672	0.036592	0.057471	0.361271
min	0.567257	0.107541	0.142414	0.555011	0.436158	0.352011	0.325507	2.284897
25%	0.608390	0.140478	0.152044	0.590531	0.467409	0.382770	0.438206	2.785926
50%	0.643161	0.178077	0.169430	0.620796	0.498937	0.396643	0.455995	3.008362
75%	0.710686	0.213784	0.180588	0.683598	0.534208	0.411103	0.477699	3.263949
max	0.722006	0.264337	0.202306	0.733828	0.604004	0.484221	0.516685	3.385636

As we can from the image, following observations can be made:

- As the Count of all the features is 10, means there is not Null values.
- The standard devaition of all the features is also very low compared to the mean, indicating that the data is not spread out(low chances of outliers).
- As i have already done the encoding of the categorical features, so no need to do it again.

5. Advanced Analysis



Insights from the correlation matrix

Focusing on this image, We can conclude :

- **avg_positve and avg_confident is highly correlated as red means they are correlated.**
- **avg_hesitant and avg_negative is also correlated as they are blue and their correlation score is 0.77(quite high) .**
- **avg_enthusiastic and avg_confident is also correlated though not as high as avg_positive and avg_confident.**
- **avg_negative and avg_confident is highly uncorrelated and their correlation score is -0.73(quite high) .**

In this way, we can see the correlation between the features like which features are dependent or which are not. From this we can if a student whose text content score is positive, then he/she is more likely more confident and enthusiastic in comparison to the student whose text content score is negative. This fact will help in further analysis.

Distribution plots were generated for various features to understand their spread and central tendencies. For example:

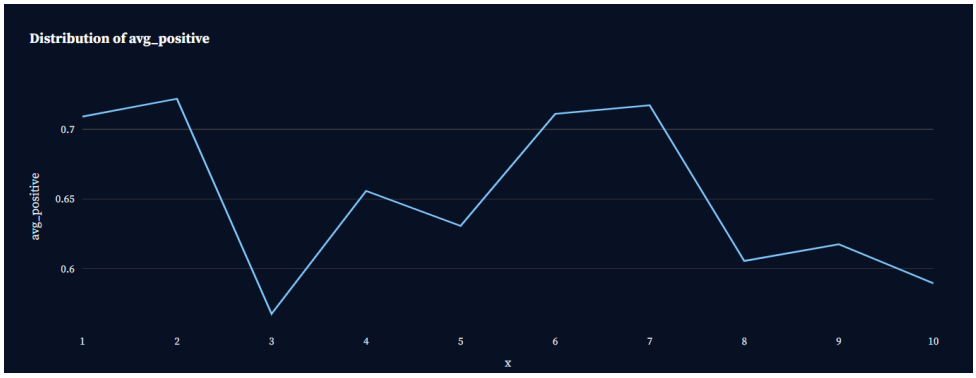


Figure 1: Distribution of avg_positive scores

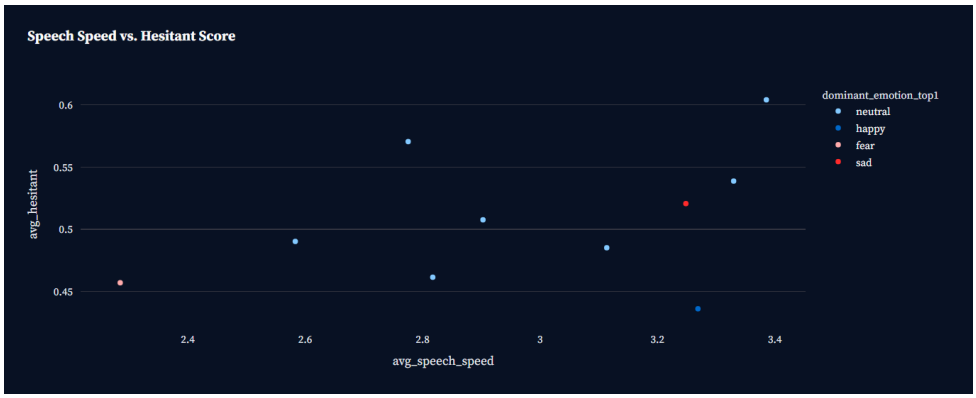
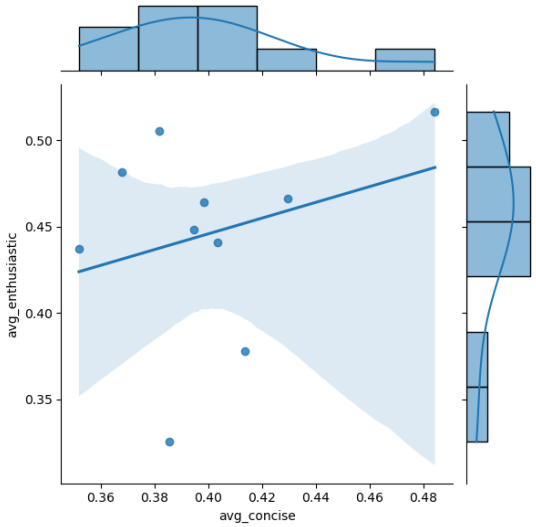


Figure 2: Speech Speed vs. Hesitant Scores

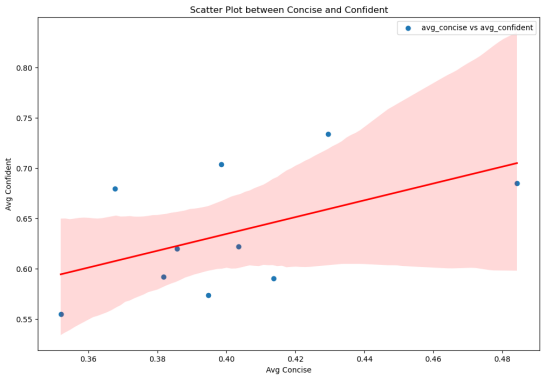
A.) Communication Skills Analysis

- **First**, I will investigate the relationship between **conciseness** and **enthusiasm** of the students.
  For this, I am plotting a joint plot between avg_concise and avg_enthusiastic. From the plot, we can see that there is a positive correlation between the two features.

  This indicates that students who are more concise in their speech are also more enthusiastic.



Joint plot between avg_enthusiastic and avg_concise.

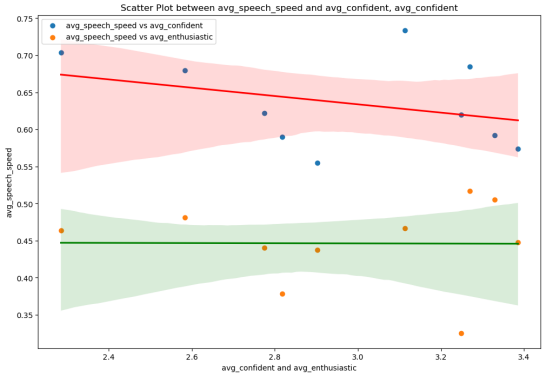


Scatter plot between avg_speech_speed and avg_hesitant.

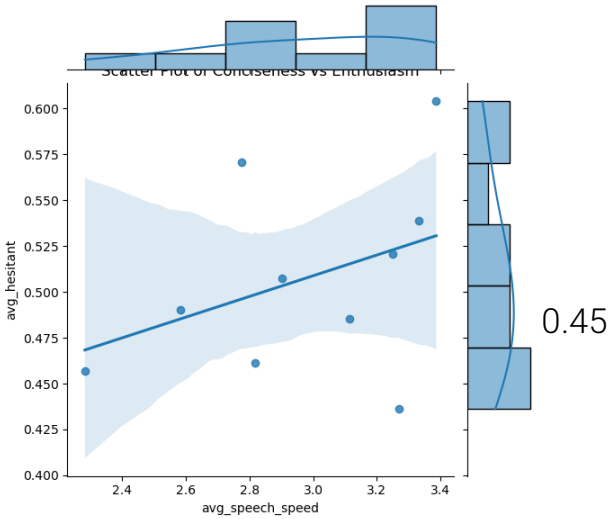
Comparison of conciseness vs enthusiasm and confidence.

- **Communication** skill is also dependent on the **speed of speech**. To analyze this, I am plotting a scatter plot between avg_speech_speed and avg_hesitant.

The plot reveals a negative correlation between these two features, meaning that students who speak faster are less hesitant in their speech.



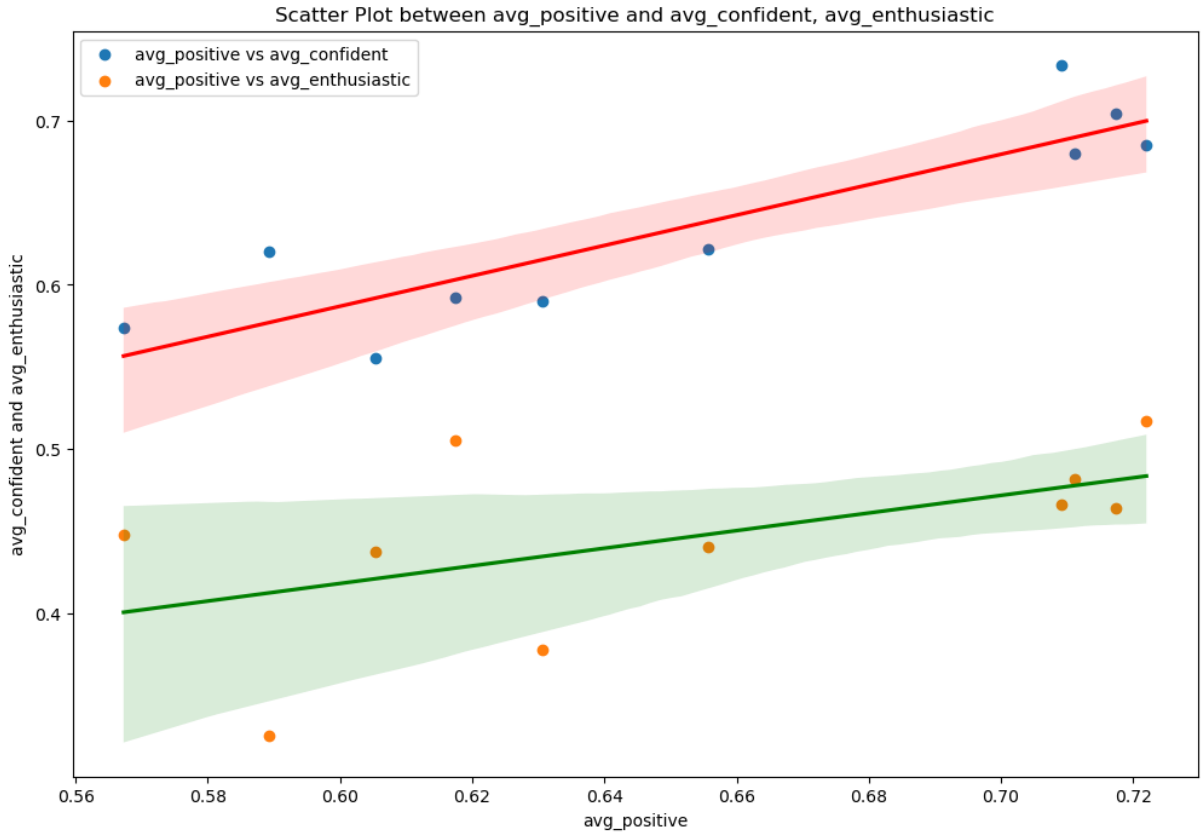
Scatter plot between avg_speech_speed and avg_confidence.



Scatter plot between avg_speech_speed and avg_hesitant.

- **Text** content scores (positive, negative, neutral) are also crucial in communication skills. Therefore, I will analyze the relationship between **positivity** and **confidence** of the students.

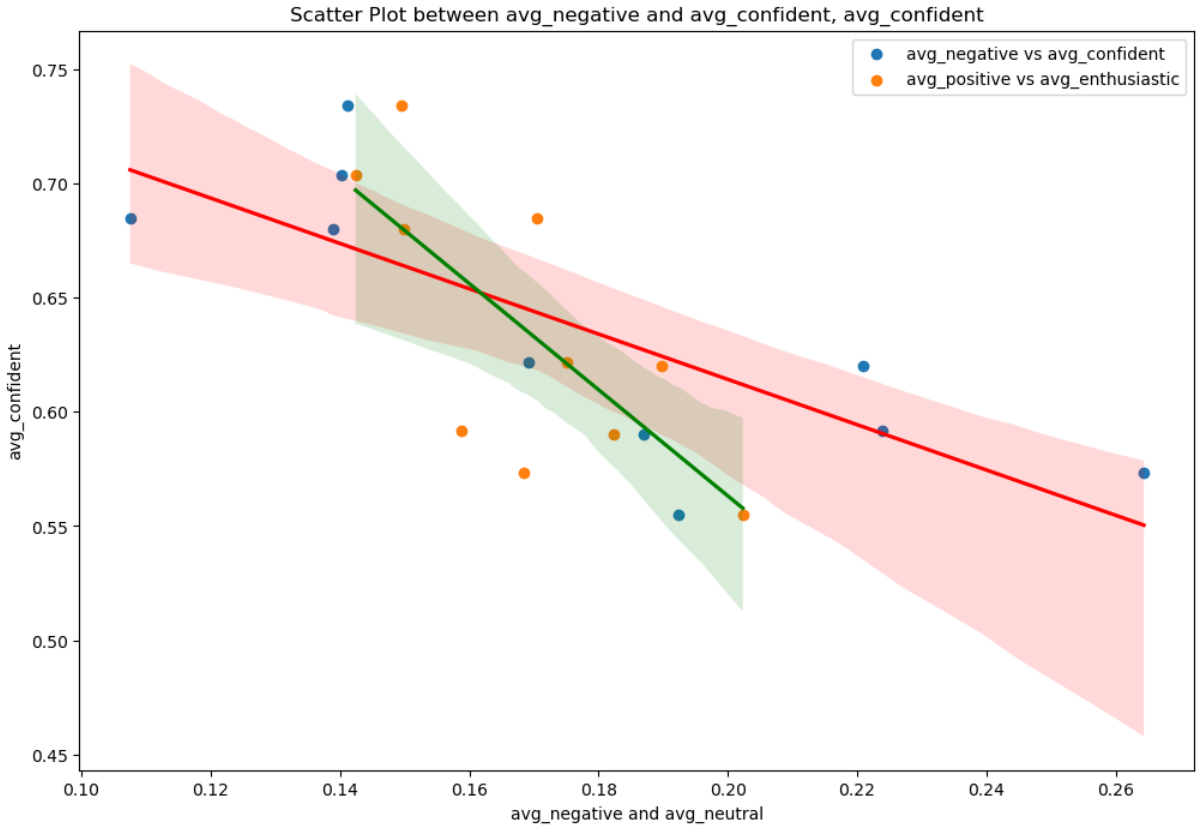
  For this, I am plotting a scatter plot between avg_positive and avg_confident. The plot shows a positive correlation between these features, indicating that students with a more positive text content score are also more confident in their speech.

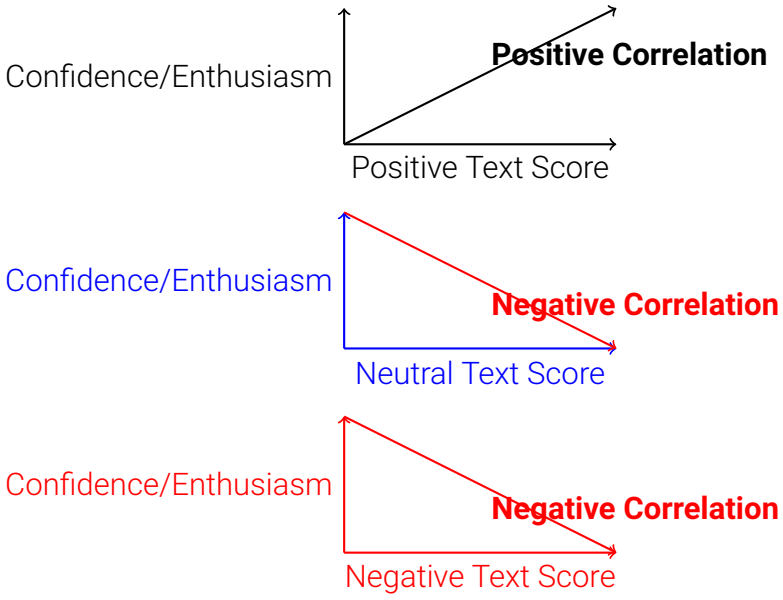


From this graph, we can clearly see the linear relation between avg_positive vs avg_confident and avg_positive vs avg_enthusiastic.

- **Additionally**, the relationship between avg_negative and avg_confident is also linear but in the opposite direction.

  This implies that students with a more negative text content score tend to be less confident and enthusiastic.





Correlation Diagram
for Text Scores with Confidence and Enthusiasm

- An interesting insight from the above graph is that avg_neutral is also linearly related to avg_confident and avg_enthusiastic, but in the opposite direction. This provides new insights into how neutrality in text content affects communication skills.

-

B.) Emotional State and Body Language Analysis

1.**Emotional Stability Analysis**
So, I have calculated the **emotional stability** of each student by analyzing the **variability** in their emotions throughout the video.

Emotional stability is a key factor in understanding how well students can manage their emotions during communication.
This is important because it helps us understand how consistent a student is in expressing different emotions.

> **Warning**
>
> Since there are 10 students and in this report, for a sample, I am showing the analysis for one student only. For other students' analyses, kindly visit the full report here.

There is slight moment in between where the student is showing fear and sadness.

This indicates that he is able to maintain a consistent emotional tone (which is neutral) during communication, which is a positive trait for effective interaction.
`This student has shown a high level of emotional stability throughout the video, with minimal fluctuations in their emotional state.`

`There is slight moment in between where the student is showing fear and sadness.`

`This indicates that he is able to maintain a consistent emotional tone (which is neutral) during communication, which is a positive trait for`

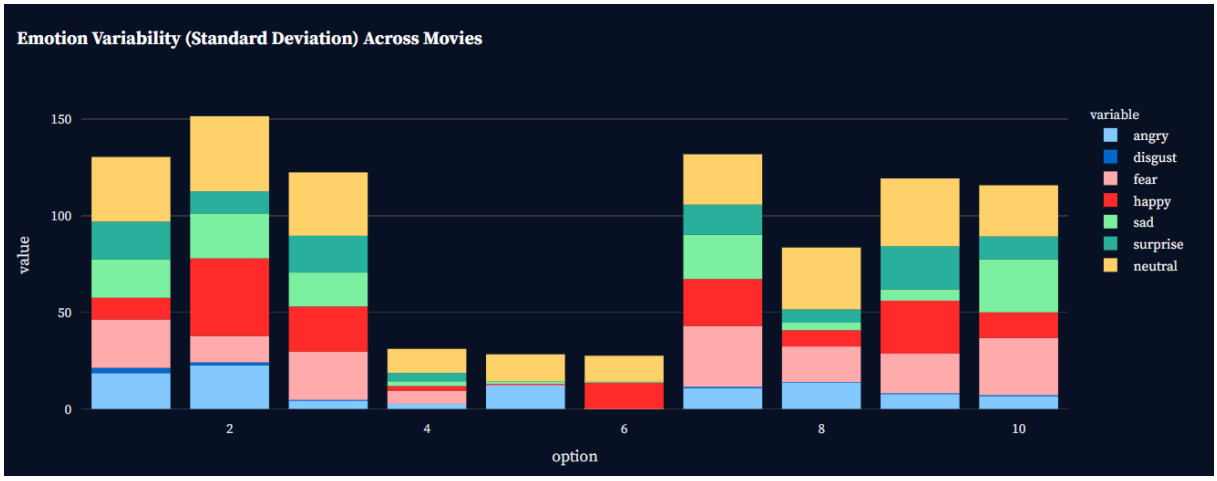
`effective interaction.`

2. **Emotion Variability Analysis**

Emotion variability is another important aspect of emotional intelligence, as it reflects how well students can adapt their emotions to different situations.

The amount of emotions a student is showing during their speech can be an indicator of their emotional intelligence.

If he shows fear or sadness for a long time, then it can be a sign of less communication skills and poor body language.



Like in this graph, we can see that except student 5 and 6, all students are showing a good amount of variability in their emotions.

3. **Body Language Analysis**
So for this, I made a new dataframe named fiNAL-GAZE_DF which contains the gaze data of all the students.

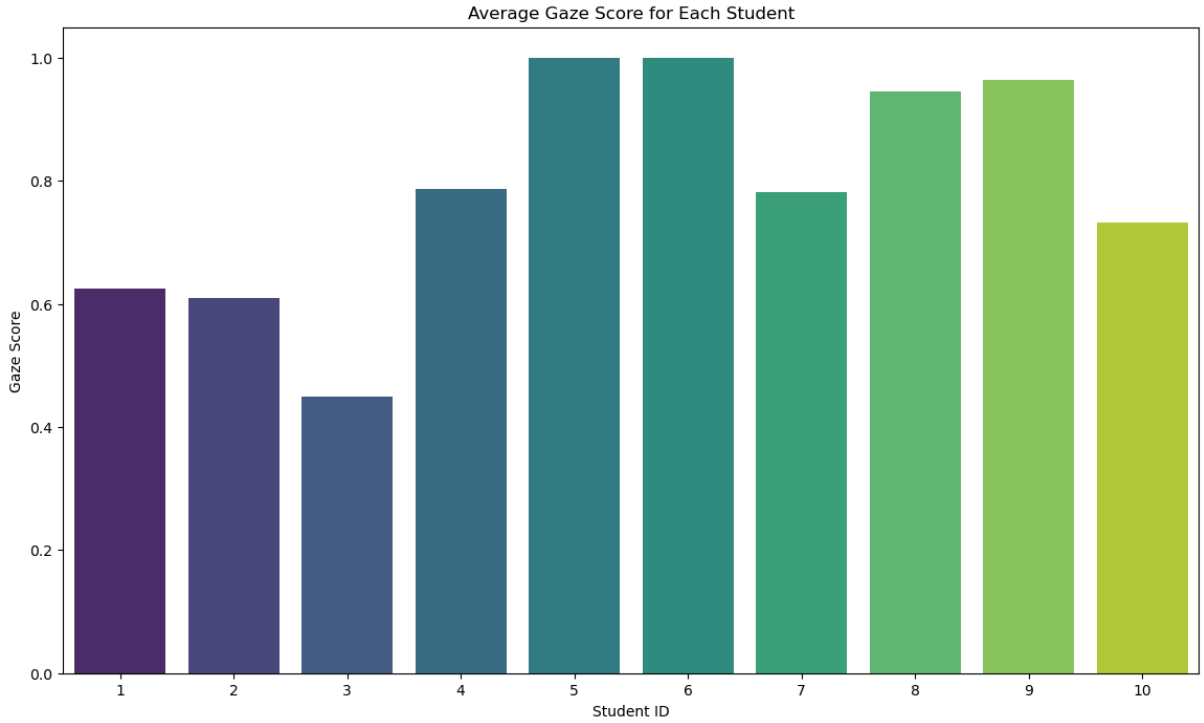
It contains the following columns:

- `movie_id:` movie_id
- `Gaze_score:` The gaze score of the student: proportion of time the candidate spends looking at the camera..
- `blink_sum:` The blink sum of the student.
- `eye_offset_std:` The eye offset standard deviation of the student.

1. If the standard deviation (std) of the eye offset of a person in a video is too high, it suggests that the person's gaze is not stable or consistent across frames

i. **Gaze Analysis**

Gaze is an important aspect of body language that can reveal a lot about a student's focus and engagement during communication.



BASED ON THE ABOVE GRAPH, FOLLOWING OBSERVATIONS CAN BE MADE:

1. Highly Engaged (0.85 - 1.0): Students who maintained frequent or constant eye contact with the camera.
   In this category, Student 5.6,8,9 fall as their gaze score is above 0.85.
2. Moderately Engaged (0.7 - 0.85): Students who maintained moderate eye contact with the camera.
   In this category, Student 4,7,10 fall as their gaze score is between 0.5 to 0.85.
3. Low Engagement (Below 0.7): Students who maintained low eye contact with the camera.
   In this category, Student 1,2,3 fall as their gaze score is below 0.5.

ii. **Blink Analysis**

Blinking is another important aspect of body language that can indicate a stu-

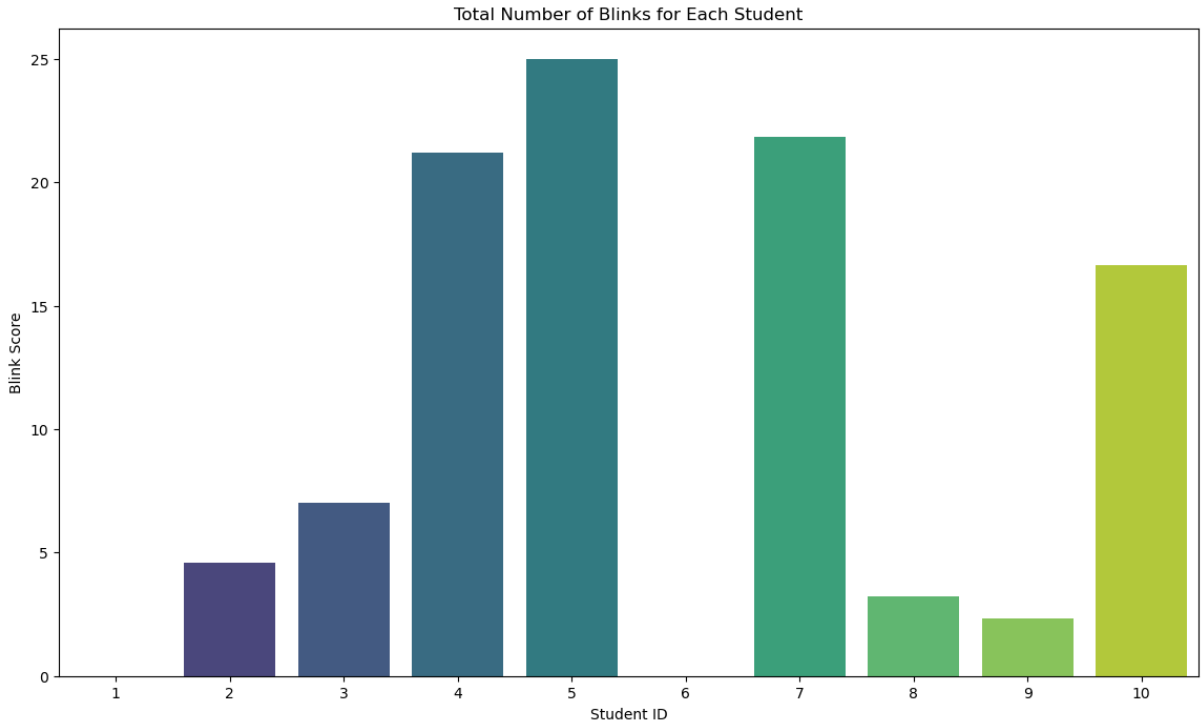
dent's level of comfort and confidence during communication.

If a student blinks too frequently, it may suggest nervousness or discomfort, while infrequent blinking may indicate confidence or focus.

Catch

Since the blinking rate will depend on the amount of time of video, i have to divided the blink$_sumwithtotal frames of the video to get the blink rate.$

Blink Rate = blink_sum / total_frames

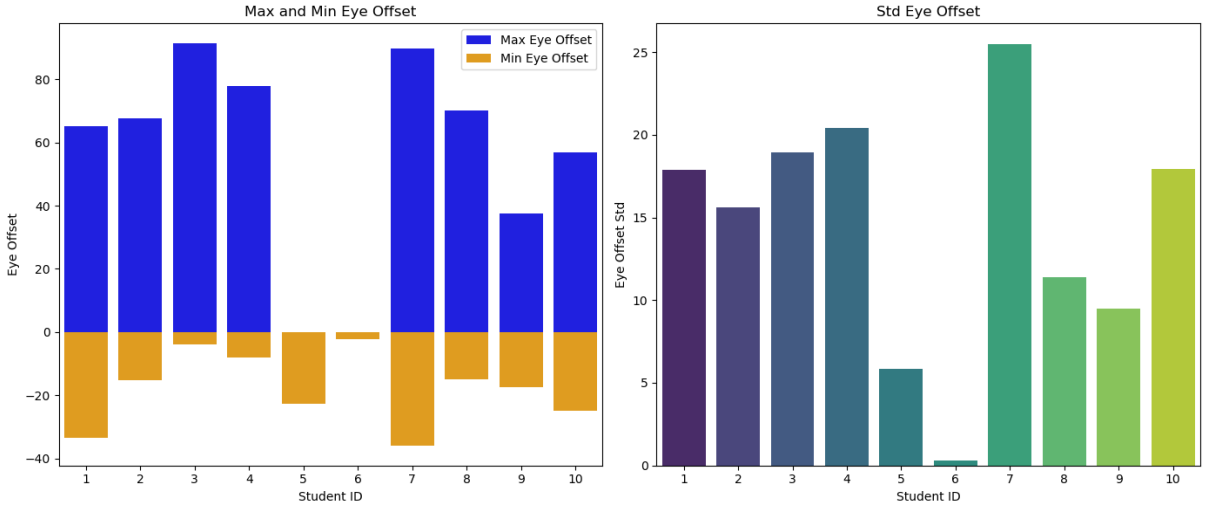


BASED ON THE ABOVE GRAPH, FOLLOWING OBSERVATIONS CAN BE MADE:

1. High Blink Rate: Students who blinked frequently during the video, i.e they are either nervour or feeling anxiety.
   In this category, Student 4,5,7,10 fall as their blink rate is .
2. Normal Blink Rate: Students who blinked at a moderate rate during the video.
   In this category, Student 1,2,3,5,6,8,9 fall as their blink rate is between 0.2 to 0.5.

iii. **Eye Offset Analysis**

Eye offset is a measure of how much a student's gaze deviates from the camera during communication.
So the higher the eye offset, the more the student's gaze is wandering away from the camera.



1. Max and Min Eye Offset (Left Plot):

This chart displays the maximum and minimum eye offsets for each student.
The blue bars represent the max positive deviation, and the orange bars represent the max negative deviation.

2. Eye Offset Standard Deviation (Right Plot):

This chart shows the standard deviation of eye offsets for each student.
The standard deviation indicates how much the student's eyes typically deviated from the mean eye position over the duration of the video.
BASED ON THE ABOVE GRAPH, FOLLOWING OBSERVATIONS CAN BE MADE:

1. `Highly Erratic Eye Movements:`
    - Student 7 shows the greatest overall deviation in both positive/negative offsets and in the standard deviation, indicating highly variable and erratic eye movements.
    - Students 3 and 4 also show significant deviation in eye movements, suggesting frequent or large fluctuations in where they were looking.
2. `Moderate Eye Movements:`
    - Students 1, 2, and 10 demonstrate moderate eye movement variability. They have noticeable deviations but are less erratic compared to students like 7 and 3.

3. `Stable Eye Movements:`

    - Students 5, 6, 8, and 9 exhibit the most stable eye movements, with minimal deviation from the mean eye position. This suggests that they maintained consistent eye contact with the camera throughout the video.

6. Candidate Evaluation Framework Explanation

This section provides a detailed explanation of the Python code used for evaluating candidates based on various metrics derived from their interview performance.

Overview

The framework consists of several key components:

- Score calculation functions for different aspects of the candidate's performance
- A function to categorize candidates based on their total score
- A main analysis function that applies the scoring framework to each candidate
- Data processing and result presentation

Score Calculation Functions

(b).1 Communication Score

$$\begin{aligned}\text{Communication Score} = \min(\text{avg\_positive} \times 10, 10) &+ \min(\text{avg\_neutral} \times 5, 5) \\ &+ \max(5 - |3 - \text{avg\_speech\_speed}| \times 2, 0) \\ &+ \text{avg\_concise} \times 5\end{aligned} \tag{1}$$

This function evaluates the candidate's communication skills based on:

- Positive and neutral language use
- Speech speed (with an ideal speed of 3)
- Conciseness

(b).2 Body Language Score

$$\text{Body Language Score} = gaze\_score * 5 + \max(5 - \frac{eye\_offset\_std}{10}, 0) + \max(5 - \frac{blink\_sum}{10}, 0) \tag{2}$$

This function assesses the candidate's body language, considering:

- Gaze direction

- Eye movement stability
- Blinking frequency

(b).3 Confidence and Enthusiasm Score

$$\text{Confidence \& Enthusiasm Score} = avg\_confident * 10 + avg\_enthusiastic * 10 \quad (3)$$

This score is a direct measure of the candidate's perceived confidence and enthusiasm.

(b).4 Emotional Intelligence Score

$$\text{Emotional Intelligence Score} = \begin{cases} 5, & \text{if dominant\_emotion\_top1 in \{happy, neutral\}} \\ 5, & \text{if dominant\_emotion\_top2 in \{happy, neutral\}} \\ -5, & \text{if dominant\_emotion\_top1 in \{sad, angry, fear\}} \\ -5, & \text{if dominant\_emotion\_top2 in \{sad, angry, fear\}} \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

This function evaluates the candidate's emotional range and appropriateness, rewarding positive emotions and penalizing negative ones.

(b).5 Composure Score

$$\text{Composure Score} = -(avg\_hesitant * 7) - (avg\_negative * 8) \quad (5)$$

This score reflects the candidate's ability to maintain composure, penalizing hesitancy and negative expressions.

Fitness Categorization

Candidates are categorized based on their total score:

$$\text{Fitness Category} = \begin{cases} \text{"Fit for role"}, & \text{if Total Score} \geq 35 \\ \text{"Okay Fit for role"}, & \text{if } 30 \leq \text{Total Score} < 35 \\ \text{"Not Fit"}, & \text{if Total Score} < 30 \end{cases} \quad (6)$$

Main Analysis Function

The `analyze_candidate` function applies all scoring functions to a candidate's data and returns a comprehensive evaluation, including:

- Individual scores for each evaluated aspect
- Total score
- Fitness category

Data Processing and Result Presentation

The framework processes the data as follows:

1. Applies the `analyze_candidate` function to each row of the input data
2. Combines the original data with the calculated results
3. Sorts candidates by their total score in descending order
4. Displays a summary table with candidate IDs, total scores, and fitness categories
5. Provides detailed feedback for each candidate, highlighting strengths and areas for improvement based on their scores in each category

Conclusion

This framework provides a comprehensive, quantitative approach to candidate evaluation, considering multiple aspects of their interview performance. It allows for objective comparison between candidates and provides actionable feedback for improvement.

7. Ranking of Students and Fit or Not

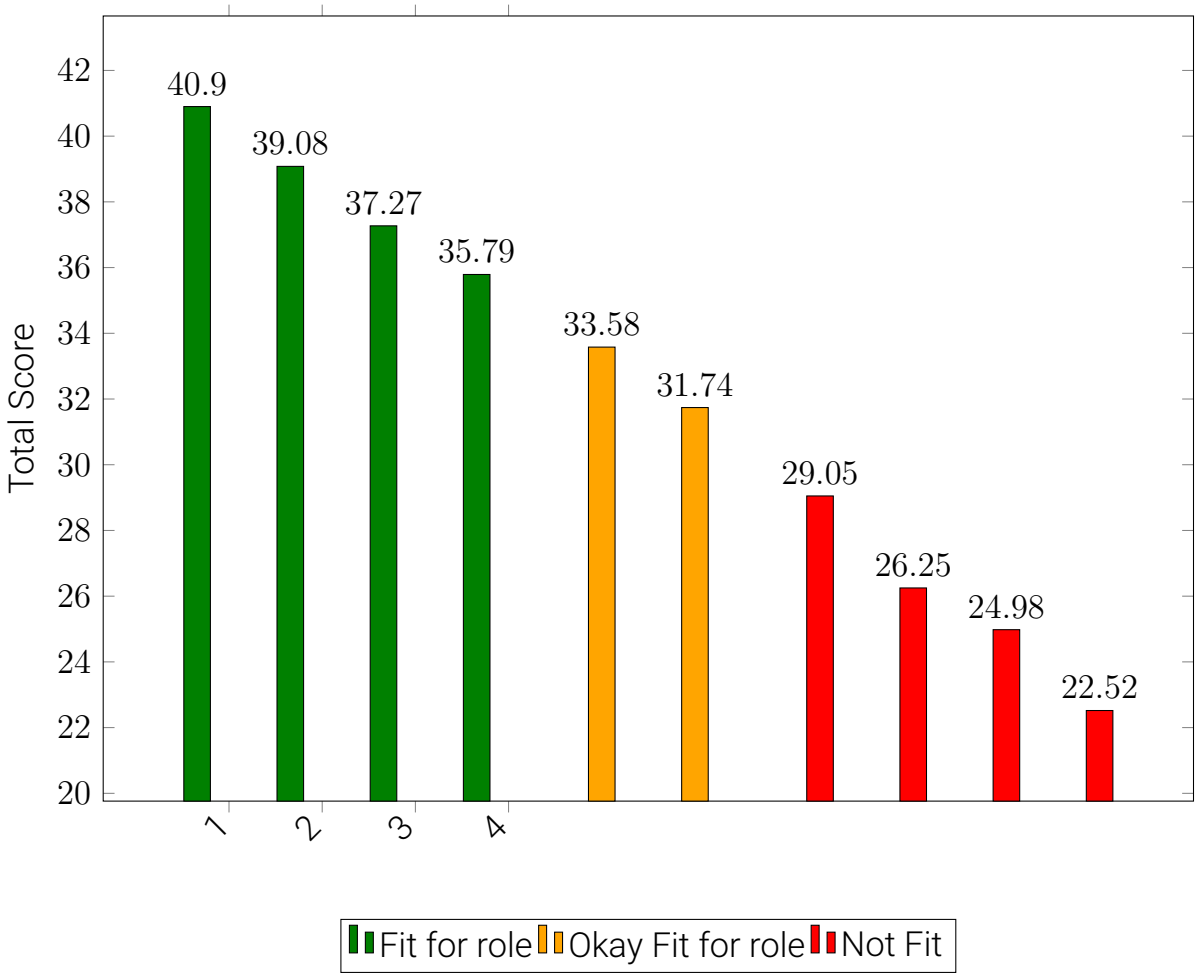


Figure 3: Candidate Total Scores

8. Candidate Evaluation Summary

Candidate ID Color Code	id	Total Score	Fitness Category
b4b6b5a2-4203-41c2-b703-c424dae1fe2b	6	40.90	Fit for role
3d7cd21a-3170-4352-b499-24ea04eaf48c	2	39.08	Fit for role
deb4a835-b82f-4f3d-b2c4-77c66eca7752	9	37.27	Fit for role
80985461-c5d6-466f-a30a-4de2784ed0a3	5	35.79	Fit for role
e2aa9258-47a5-46ab-9c5c-283460f7a807	1	33.58	Okay for role
1c0c686b-3aae-4ac6-8625-3e86a7a0892f	8	31.74	Okay for role
62ea9b36-7860-4dc9-827c-600604286571	4	29.05	Not Fit
f299e1b2-7d92-4420-9c5a-d0d2590abdbe	3	26.25	Not Fit
d851fe95-3ead-47c1-88aa-d6fc453f7021	7	24.98	Not Fit
70a013ed-120a-41fa-bedd-75a5d15afb76	10	22.52	Not Fit

Table 1: Candidate Evaluation Overview

().1 ROW ID: 6

Candidate ID: b4b6b5a2-4203-41c2-b703-c424dae1fe2b

Movie ID: 92016995-e455-4651-9f6e-fbca0d423f21

- **Total Score:** 40.90
- **Fitness Category:** Fit for role
- **Areas for Improvement:**
  - Enhance communication skills

  - Work on body language
  - Boost confidence and enthusiasm
  - Develop emotional intelligence
  - Improve composure under pressure

().2 ROW ID: 2

Candidate ID: 3d7cd21a-3170-4352-b499-24ea04eaf48c

Movie ID: baa26895-85b2-465b-a972-649b41d9870e

- **Total Score:** 39.08
- **Fitness Category:** Fit for role
- **Areas for Improvement:**
  - Enhance communication skills
  - Work on body language
  - Boost confidence and enthusiasm
  - Develop emotional intelligence
  - Improve composure under pressure

().3 ROW ID: 9

Candidate ID: deb4a835-b82f-4f3d-b2c4-77c66eca7752

Movie ID: dfb0d746-609f-4dac-8e1d-c0325fb64394

- **Total Score:** 37.27
- **Fitness Category:** Fit for role
- **Areas for Improvement:**
  - Enhance communication skills
  - Work on body language
  - Boost confidence and enthusiasm
  - Develop emotional intelligence
  - Improve composure under pressure

().4 ROW ID: 5

Candidate ID: 80985461-c5d6-466f-a30a-4de2784ed0a3

Movie ID:9c350343-e895-49df-af90-d50b91d19d3e

- **Total Score:** 35.79
- **Fitness Category:** Fit for role
- **Areas for Improvement:**
  - Enhance communication skills
  - Work on body language
  - Boost confidence and enthusiasm
  - Develop emotional intelligence
  - Improve composure under pressure

Okay Fit for Role Candidates

(a).1 Row ID: 1

Candidate ID: e2aa9258-47a5-46ab-9c5c-283460f7a807

Movie ID: 93663f94-bf0a-4ce8-a29a-a5236cc7fe6a

- **Total Score:** 33.58
- **Fitness Category:** Okay Fit for role
- **Areas for Improvement:**
  - Enhance communication skills
  - Work on body language
  - Boost confidence and enthusiasm
  - Develop emotional intelligence
  - Improve composure under pressure

(a).2 Row ID: 8

Candidate ID: 1c0c686b-3aae-4ac6-8625-3e86a7a0892f

Movie ID: 813af424-a584-4417-b7ee-0d4c705e83c9

- **Total Score:** 31.74

- **Fitness Category:** Okay Fit for role
- **Areas for Improvement:**
  - Enhance communication skills
  - Work on body language
  - Boost confidence and enthusiasm
  - Develop emotional intelligence
  - Improve composure under pressure

Not Fit Candidates

(b).1 Row ID: 4

Candidate ID: 62ea9b36-7860-4dc9-827c-600604286571

Movie ID: 6b0386fc-41de-4196-b0d6-3d0b815c2dbc

- **Total Score:** 29.05
- **Fitness Category:** Not Fit
- **Areas for Improvement:**
  - Enhance communication skills
  - Work on body language
  - Boost confidence and enthusiasm
  - Develop emotional intelligence
  - Improve composure under pressure

(b).2 Row ID: 3

Candidate ID: f299e1b2-7d92-4420-9c5a-d0d2590abdbe

Movie ID: d0b9170b-98b9-48e1-a1b2-1d661bb0d853

- **Total Score:** 26.25
- **Fitness Category:** Not Fit
- **Areas for Improvement:**
  - Enhance communication skills
  - Work on body language
  - Boost confidence and enthusiasm
  - Develop emotional intelligence
  - Improve composure under pressure

(b).3 Row ID: 7

Candidate ID: d851fe95-3ead-47c1-88aa-d6fc453f7021

Movie ID: 6539370c-256e-4ed2-9d00-1be1f051163f

- **Total Score:** 24.98
- **Fitness Category:** Not Fit
- **Areas for Improvement:**
  - Enhance communication skills
  - Work on body language
  - Boost confidence and enthusiasm
  - Develop emotional intelligence
  - Improve composure under pressure

(b).4 Row ID: 10

Candidate ID: 70a013ed-120a-41fa-bedd-75a5d15afb76

Movie ID: 83c20b83-7881-499d-a40d-cc06b65869f8

- **Total Score:** 22.52
- **Fitness Category:** Not Fit
- **Areas for Improvement:**
  - Enhance communication skills
  - Work on body language
  - Boost confidence and enthusiasm
  - Develop emotional intelligence
  - Improve composure under pressure

9.Detailed Individual Student Analysis

On the analysis i did in section advanced analysis, I did seperate analysis for each student.
on his expertise, communication skills, emotional intelligence, and body language are analyzed in detail.

Transcript Data (Positive, Negative, Neutral):
Higher positive score indicates better ability to maintain a positive tone, which is critical for many roles. Higher negative score may indicate a candidate who tends to express negative emotions or perspectives, which may reduce their fit. A balance of neutral responses is normal, but too much neutrality may indicate a lack of enthusiasm or engagement.
Confidence:
Higher confidence indicates a candidate who is more assured in their communication, a key factor in leadership or client-facing roles.
Hesitance:
Higher hesitance suggests uncertainty, which may negatively impact the fit for leadership or critical decision-making roles. However, it can be okay for technical or supporting roles.
Speech Speed:
Optimal speech speed (not too fast or too slow) shows clarity in communication. Extremes in either direction can indicate nervousness or a lack of engagement.
Enthusiasm:
Candidates who show higher enthusiasm are more likely to be motivated and engaging, which is important in most roles.
Body Language (Blink Sum, Gaze Score, Eye Offset):

- Blink Sum: Too many blinks may indicate stress or nervousness.
- Gaze Score: Higher gaze scores show good eye contact, which is important for building rapport.

- Eye Offset: Large deviations might indicate discomfort or lack of focus.

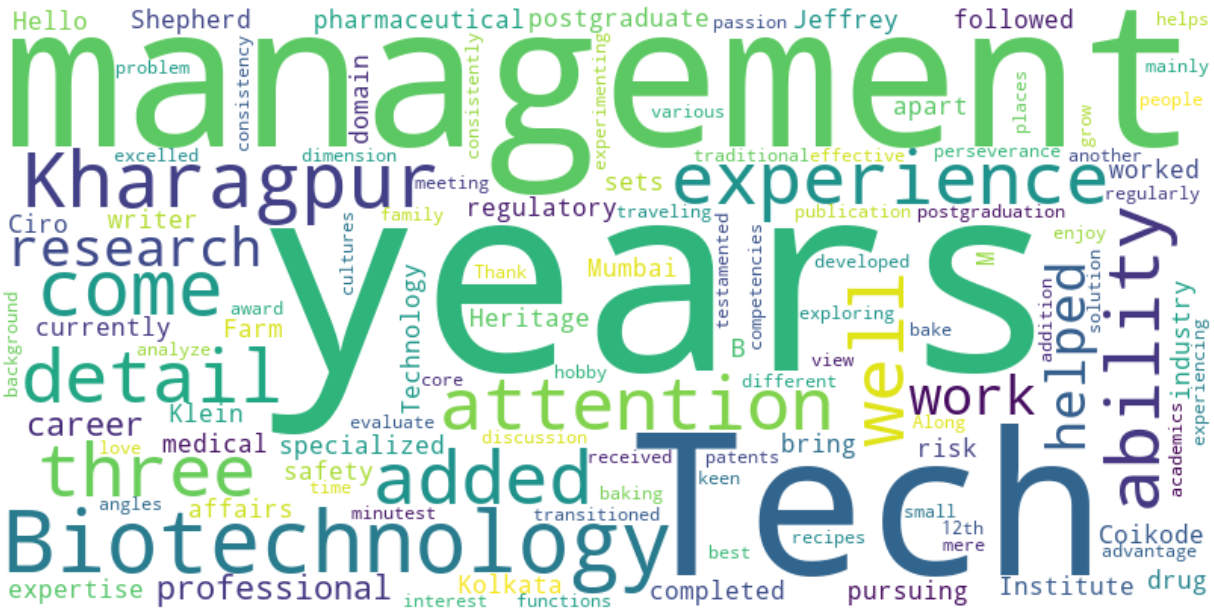
Student 1

Transcript Sentiment:

Sentiment Breakdown

1. **Positive Score:** 0.709199
2. **Negative Score:** 0.141214
3. **Neutral Score:** 0.149586

Analysis: The student's speech is predominantly **positive**, with minimal negative or neutral tones.



Area of Expertise:

- Expertise in biotechnology, management, and regulatory affairs, with a focus on pharmaceutical industry experience.
- Three years of work in medical writing, drug safety, and research, with attention to detail and award-winning research background.

Confidence and Hesitation:

Confidence and Hesitant Scores

1. **Confidence Score:** 0.733828
2. **Hesitant Score:** 0.485172

Analysis: The student demonstrates high **confidence**, although there is some hesitation present.

- **Strengths:** Strong confidence helps in engaging the audience and conveying ideas effectively.
- **Weaknesses:** Moderate hesitation might reduce the clarity of their message.

Enthusiasm:

Enthusiastic Score

1. **Enthusiastic Score:** 0.466497

Analysis: The student has a high enthusiasm score, indicating a positive energy in their speech.

- **Communication Skills:** High enthusiasm combined with confidence makes their speech engaging and dynamic.

Speech Speed:

Speech Speed

1. **Speech Speed:** 0.416344

Analysis: The student's speech speed is close to normal speed, indicating the stability and .

- **Emotional Intelligence:** Their ability to maintain a good pace shows awareness of audience engagement.

Dominant Emotions:

Emotional State

1. **Top 1 Emotion:** Neutral
2. **Top 2 Emotion:** Fear

Analysis: The student's emotional state is predominantly neutral, with some underlying fear present.

- **Emotional Stability:** The neutral tone suggests good emotional control, although the presence of fear may indicate some anxiety.

Student 1: Comparative Analysis

- **Strengths:**
  - **Confidence Score:** Highest among all students, indicating strong self-assurance and control during speech (0.733828).
  - **Positive Sentiment:** 2nd highest positive score (0.709199), indicating a mostly optimistic and upbeat tone.
  - **Speech Speed:** Moderately slow pace (0.416344), which helps in clarity and audience comprehension.
- **Weaknesses:**
  - **Hesitation Score:** Slightly high (0.485172), suggesting occasional pauses or uncertainty in speech delivery.
  - **Enthusiasm Score:** On the lower side (0.466497), which may reduce audience engagement and energy.
- **Areas of Expertise:**
  - **Confidence:** Excellent control over speech, useful in leadership and formal presentations.
  - **Clear Communication:** Balanced speech pace with minimal rush, ensuring clarity of ideas.

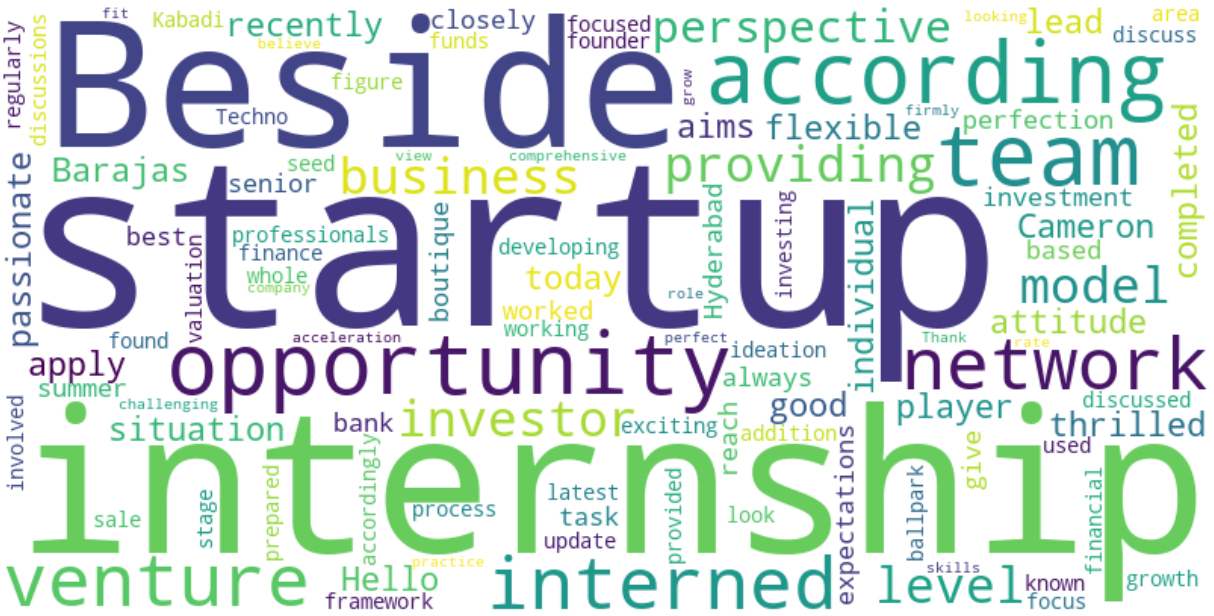
Student 2

Transcript Sentiment:

Sentiment Breakdown

1. **Positive Score:** 0.722006
2. **Negative Score:** 0.107541
3. **Neutral Score:** 0.170453

Area of Expertise based on the wordblob image:



Area of Expertise:

- Specialized in startups, internships, and venture creation, with a focus on providing opportunities and team dynamics.
- Experience in investor relations, business model development, and working with seed-stage ventures.

Confidence and Hesitation:

Confidence and Hesitant Scores

1. **Confidence Score:** 0.684879
2. **Hesitant Score:** 0.436158

Analysis: The student demonstrates high **confidence**(second highest score), although there is very few hesitation present(lowest). **Enthusiasm:**

Enthusiastic Score

1. **Enthusiastic Score:** 0.516685

Analysis: The student has a highest enthusiasm score among all the students, indicating a positive energy in their speech. **Speech Speed:**

Speech Speed

1. **Speech Speed:** 0.869530

Analysis: The student's speech speed is fastest among all the students, likely due to their enthusiasm and confidence. **Dominant Emotions:**

Emotional State

1. **Top 1 Emotion:** Happy
2. **Top 2 Emotion:** Neutral

Student 2: Comparative Analysis

- **Strengths:**
  - **Positive Sentiment:** Highest among all students (0.722006), reflecting a highly positive and enthusiastic tone.
  - **Enthusiasm Score:** Highest among the group (0.516685), showing an energetic and engaging delivery.

  - **Confidence Score:** 3rd highest (0.684879), indicating a strong but slightly lower confidence than Student 1.

- **Weaknesses:**
  - **Speech Speed:** Fastest among all students (0.869530), which might make it difficult for the audience to follow.
  - **Hesitation Score:** Moderate (0.436158), showing occasional signs of uncertainty.
- **Areas of Expertise:**
  - **Enthusiasm:** High energy levels make Student 2's speech engaging, useful in persuasive or motivational settings.
  - **Positive Outlook:** The optimistic tone is effective in inspiring confidence and connection with the audience.

Analysis: The student's emotional state is predominantly happy, with some underlying neutral present.

Student 3

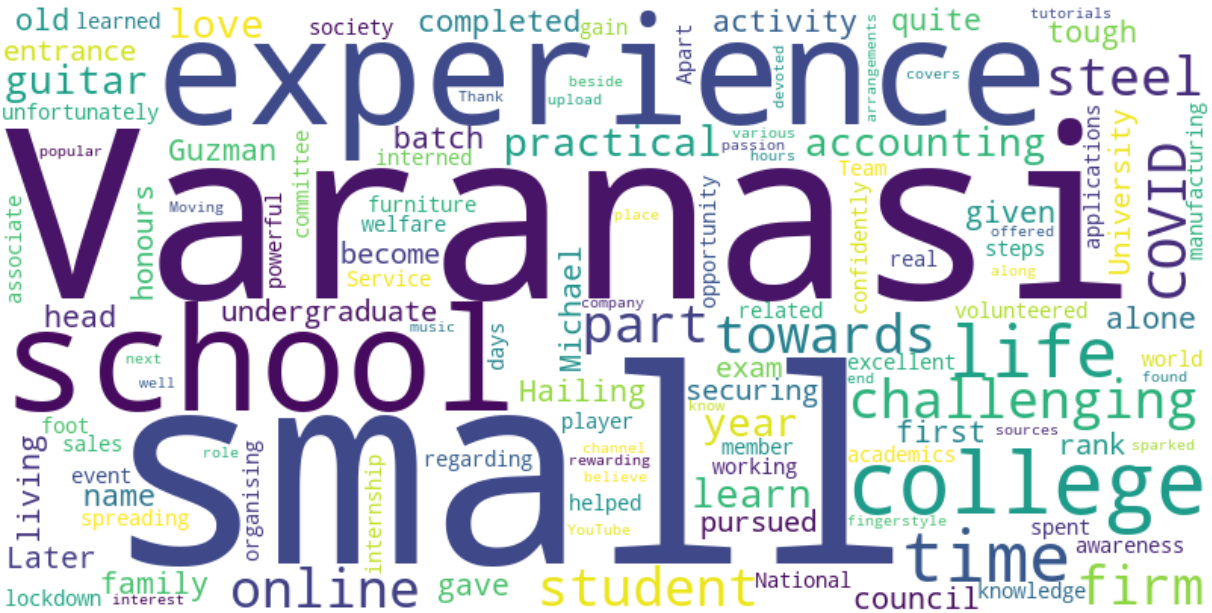
Transcript Sentiment:

Sentiment Breakdown

1. **Positive Score:** 0.652301 (Rank: 7/10)
2. **Negative Score:** 0.167569 (Rank: 7/10)
3. **Neutral Score:** 0.180130 (Rank: 6/10)

Analysis: Balanced sentiment with moderate positive and negative scores, indicating a generally neutral emotional stance with a slight positive tilt.

Area of Expertise based on the wordblob image:



Area of Expertise based on the wordblob image:

- The candidate has significant experience in education and student life.
- They possess practical skills and knowledge in areas like accounting, steel manufacturing, and internships, with a focus on real-world applications and challenging experiences.
- Additional interests include guitar, leadership roles, and online learning, suggesting a well-rounded background with both academic and extracurricular engagement.

Confidence and Hesitation:

Confidence and Hesitant Scores

1. **Confidence Score:** 0.663123 (Rank: 4/10)
2. **Hesitant Score:** 0.445621 (Rank: 9/10)

Analysis: High confidence with relatively high hesitation. This indicates strong self-assurance but also a significant level of nervousness. **Enthusiasm:**

Enthusiastic Score

1. **Enthusiastic Score:** 0.418567 (Rank: 10/10)

Analysis: Lowest enthusiasm score, suggesting lower engagement and energy compared to peers. **Speech Speed:**

Speech Speed

1. **Speech Speed:** -0.311245 (Rank: 6/10)

Analysis: Moderate speech speed, indicating a balanced delivery pace. **Dominant Emotions:**

Emotional State

1. **Top 1 Emotion:** Neutral
2. **Top 2 Emotion:** Fear

Analysis: Predominantly neutral with fear as a secondary emotion, reflecting a mostly calm demeanor with some underlying apprehension. **Comparative Analysis:**

- **Strengths:**
  - **Confidence Score:** 4th highest among all students, showing strong self-assurance.
  - **Speech Speed:** 6th in ranking, indicating a moderate pace in speaking.
- **Weaknesses:**
  - **Enthusiasm Score:** Lowest enthusiasm, indicating reduced engagement and energy.
  - **Hesitant Score:** 9th highest, reflecting considerable hesitation compared to peers.
- **Areas of Expertise:**
  - **Confidence:** Demonstrates a high level of confidence, useful in scenarios requiring strong self-belief.
  - **Speech Pace:** Balanced speech speed, indicating a moderate approach to communication.

Overall: Student 3 shows strong confidence but needs to work on increasing enthusiasm and reducing hesitation to improve overall presentation and engagement.

Student 4

Transcript Sentiment:

Sentiment Breakdown

1. **Positive Score:** 0.655748 (Rank: 4/10)
2. **Negative Score:** 0.169142 (Rank: 6/10)
3. **Neutral Score:** 0.175110 (Rank: 6/10)

Analysis: The sentiment is slightly positive with a well-balanced neutral and negative score, indicating a generally positive outlook with some neutral stance.

Area of Expertise based on the wordblob image:

- Academic background in engineering, with a focus on data analysis and machine learning
- Experience in Python programming and various technical projects, possibly including VLSI and electronics
- Involvement in academic administration and leadership roles, potentially at a center or school level

Confidence and Hesitation:

Confidence and Hesitant Scores

1. **Confidence Score:** 0.621740 (Rank: 5/10)
2. **Hesitant Score:** 0.570452 (Rank: 6/10)

Analysis: Moderate confidence with high hesitation, showing a good level of self-assurance but also significant nervousness. **Enthusiasm:**

Enthusiastic Score

1. **Enthusiastic Score:** 0.440626 (Rank: 9/10)

Analysis: Low enthusiasm, indicating reduced engagement and energy compared to peers. **Speech Speed:**

Speech Speed

1. **Speech Speed:** -0.570773 (Rank: 7/10)

Analysis: Slow speech speed, suggesting a deliberate and possibly cautious delivery pace. **Dominant Emotions:**

Emotional State

1. **Top 1 Emotion:** Neutral
2. **Top 2 Emotion:** Fear

Analysis: Predominantly neutral with a secondary emotion of fear, reflecting a generally calm demeanor with some apprehension. **Comparative Analysis:**

- **Strengths:**
  - **Sentiment:** 4th in positive score, showing a good positive outlook.
  - **Confidence:** 5th highest, indicating solid self-assurance.
- **Weaknesses:**
  - **Enthusiasm Score:** 9th lowest, suggesting lower engagement.
  - **Speech Speed:** 7th slowest, reflecting a slower delivery pace.
  - **Hesitant Score:** 6th highest, showing notable hesitation.
- **Areas of Expertise:**
  - **Sentiment:** Positive sentiment indicates a generally favorable attitude.
  - **Confidence:** Good level of self-assurance in communication.

Overall: Student 4 demonstrates solid confidence and positive sentiment but should work on increasing enthusiasm and reducing hesitation to enhance overall delivery.

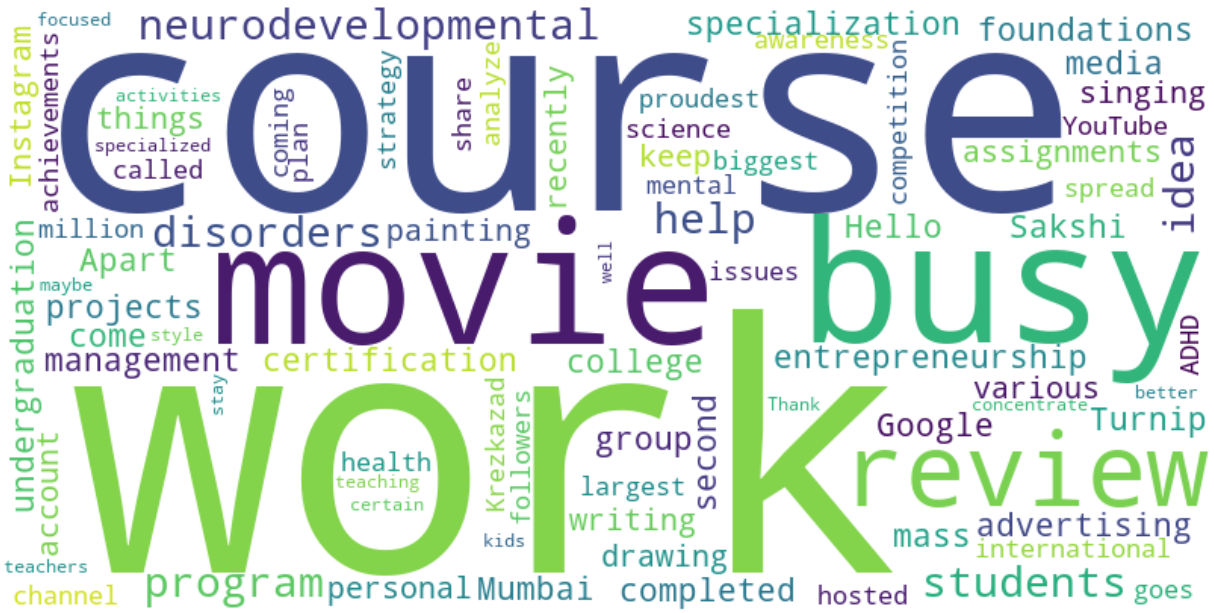
Student 5

Transcript Sentiment:

Sentiment Breakdown

1. **Positive Score:** 0.630573 (Rank: 6/10)
2. **Negative Score:** 0.187013 (Rank: 6/10)
3. **Neutral Score:** 0.182414 (Rank: 5/10)

Analysis: The sentiment shows a positive outlook with a balanced negative and neutral stance, indicating a generally optimistic but somewhat cautious approach. **Area of Expertise based on the wordblob image:**



- Expertise in neurodevelopmental courses and disorders
- Experience in movie reviews and critique
- Involvement in educational programs, possibly including entrepreneurship and project management

Confidence and Hesitation:

Confidence and Hesitant Scores

1. **Confidence Score:** 0.590094 (Rank: 7/10)
2. **Hesitant Score:** 0.461488 (Rank: 8/10)

Analysis: Moderate confidence with high hesitation, reflecting a solid level of self-assurance but with considerable nervousness. **Enthusiasm:**

Enthusiastic Score

1. **Enthusiastic Score:** 0.378110 (Rank: 10/10)

Analysis: Lowest enthusiasm score, indicating the least amount of engagement and energy compared to peers. **Speech Speed:**

Speech Speed

1. **Speech Speed:** -0.448559 (Rank: 8/10)

Analysis: Slow speech speed, indicating a cautious or deliberate speaking pace. **Dominant Emotions:**

Emotional State

1. **Top 1 Emotion:** Neutral
2. **Top 2 Emotion:** Neutral

Analysis: Predominantly neutral with no strong secondary emotion, reflecting a calm and even emotional state. **Comparative Analysis:**

- **Strengths:**
  - **Sentiment:** 6th in positive score, showing a favorable outlook.
  - **Confidence:** 7th highest, indicating a reasonable level of self-assurance.
- **Weaknesses:**
  - **Enthusiasm Score:** 10th lowest, showing the least energy and engagement.
  - **Speech Speed:** 8th slowest, suggesting a slower delivery pace.
  - **Hesitant Score:** 8th highest, indicating significant hesitation.
- **Areas of Expertise:**
  - **Sentiment:** Positive outlook with balanced emotional stance.

– **Confidence:** Good self-assurance in communication.

Overall: Student 5 has a favorable sentiment and reasonable confidence but should focus on increasing enthusiasm and reducing hesitation to improve engagement and delivery.

Student 6

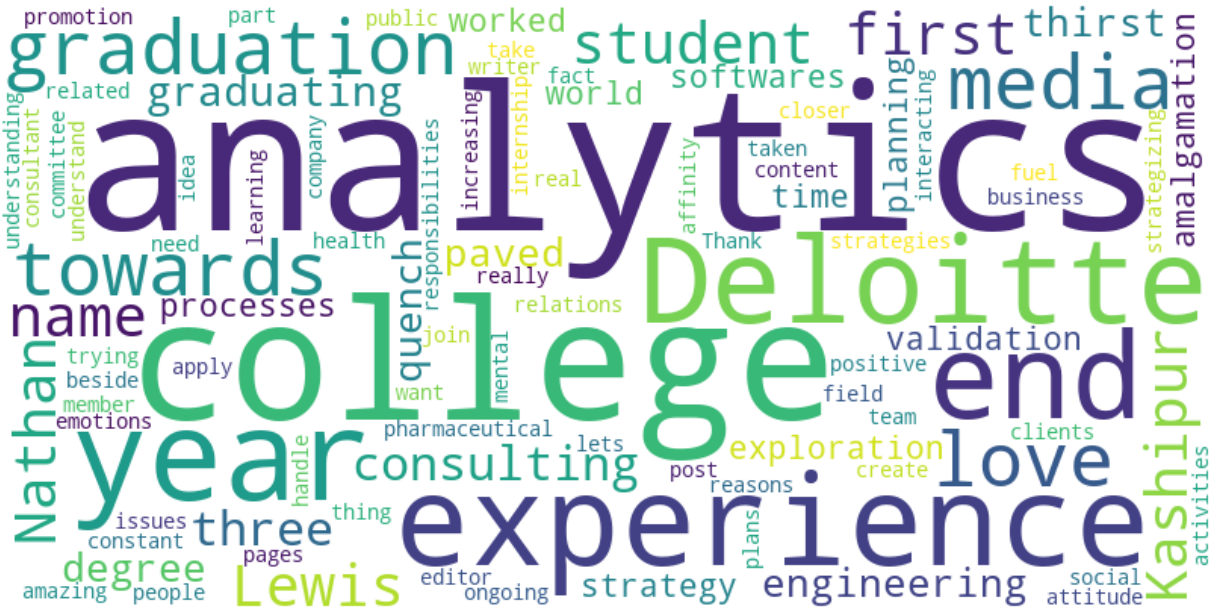
Transcript Sentiment:

Sentiment Breakdown

1. **Positive Score:** 0.711182 (Rank: 2/10)
2. **Negative Score:** 0.138992 (Rank: 8/10)
3. **Neutral Score:** 0.149826 (Rank: 8/10)

Analysis: Strongly positive sentiment with low negative and neutral scores, indicating an optimistic and positive outlook. **Area of**

Expertise based on the wordblob image:



- Analytics experience, particularly with Deloitte

- Focus on media and business consulting
- Background in college education and graduation-related activities

Confidence and Hesitation:

Confidence and Hesitant Scores

1. **Confidence Score:** 0.679755 (Rank: 3/10)
2. **Hesitant Score:** 0.490252 (Rank: 7/10)

Analysis: High confidence with moderate hesitation, showing strong self-assurance but with some nervousness. **Enthusiasm:**

Enthusiastic Score

1. **Enthusiastic Score:** 0.481433 (Rank: 6/10)

Analysis: Moderate enthusiasm, reflecting a balanced level of engagement and energy. **Speech Speed:**

Speech Speed

1. **Speech Speed:** -1.131826 (Rank: 9/10)

Analysis: Slow speech speed, indicating a measured and deliberate delivery. **Dominant Emotions:**

Emotional State

1. **Top 1 Emotion:** Neutral
2. **Top 2 Emotion:** Neutral

Analysis: Predominantly neutral with a consistent emotional state. **Comparative Analysis:**

- **Strengths:**
  - **Sentiment:** 2nd in positive score, showing a very positive outlook.

  - **Confidence:** 3rd highest, reflecting strong self-assurance.

- **Weaknesses:**
  - **Speech Speed:** 9th slowest, suggesting a slower speaking pace.
  - **Hesitant Score:** 7th highest, indicating moderate hesitation.
- **Areas of Expertise:**
  - **Sentiment:** Highly positive outlook.
  - **Confidence:** Strong self-assurance in communication.

Overall: Student 6 exhibits a very positive sentiment and strong confidence but should work on managing their speech speed and hesitation to enhance overall communication effectiveness.

Student 7

Transcript Sentiment:

Sentiment Breakdown

1. **Positive Score:** 0.717354 (Rank: 1/10)
2. **Negative Score:** 0.140232 (Rank: 9/10)
3. **Neutral Score:** 0.142414 (Rank: 9/10)

Analysis: Extremely positive sentiment with minimal negative and neutral scores, indicating an exceptionally optimistic and upbeat outlook. **Area of Expertise based on the wordblob image:**

- Background in heritage and cultural studies
- Skills related to insurance and risk management
- Possible academic experience, likely at a university (potentially Banaras Hindu University)

Confidence and Hesitation:

Confidence and Hesitant Scores

1. **Confidence Score:** 0.703714 (Rank: 2/10)
2. **Hesitant Score:** 0.457070 (Rank: 5/10)

Analysis: High confidence with lower hesitation, showing strong self-assurance and less nervousness. **Enthusiasm:**

Enthusiastic Score

1. **Enthusiastic Score:** 0.463940 (Rank: 5/10)

Analysis: Moderate enthusiasm, indicating a good level of engagement and energy. **Speech Speed:**

Speech Speed

1. **Speech Speed:** -2.002086 (Rank: 10/10)

Analysis: Slowest speech speed, reflecting a very deliberate and cautious delivery. **Dominant Emotions:**

Emotional State

1. **Top 1 Emotion:** Fear
2. **Top 2 Emotion:** Sad

Analysis: Predominantly fearful with some sadness, indicating emotional challenges despite a positive sentiment. **Comparative Analysis:**

- **Strengths:**
  - **Sentiment:** 1st in positive score, demonstrating an exceptionally optimistic outlook.
  - **Confidence:** 2nd highest, reflecting strong self-assurance.
- **Weaknesses:**
  - **Speech Speed:** 10th slowest, suggesting a very slow speaking pace.
  - **Emotional State:** Predominantly fearful and sad, indicating underlying emotional challenges.
- **Areas of Expertise:**
  - **Sentiment:** Extremely positive and optimistic outlook.
  - **Confidence:** Strong self-assurance in communication.

Overall: Student 7 exhibits exceptional positivity and strong confidence but should address the slow speech pace and manage underlying emotional challenges to enhance overall effectiveness.

Student 8

Transcript Sentiment:

Sentiment Breakdown

1. **Positive Score:** 0.605402 (Rank: 5/10)
2. **Negative Score:** 0.192292 (Rank: 5/10)
3. **Neutral Score:** 0.202306 (Rank: 4/10)

Analysis: Balanced sentiment with a positive outlook and moderate neutral and negative scores, indicating a generally positive but cautious approach. **Area of Expertise based on the wordblob image:**

- **Audit and Accounting**: Expertise in statutory audit, chartered accountancy, and internal audit practices.
- **Data and Analytics**: Proficiency in interpreting and working with statistical data, financial analysis, and analytical tools.
- **Programming and Coding**: Experience in coding and developing solutions, possibly related to automation or data processing.

Confidence and Hesitation:

Confidence and Hesitant Scores

1. **Confidence Score:** 0.555011 (Rank: 8/10)
2. **Hesitant Score:** 0.507622 (Rank: 4/10)

Analysis: Moderate confidence with lower hesitation, showing a reasonable level of self-assurance with some nervousness.

Enthusiasm:

Enthusiastic Score

1. **Enthusiastic Score:** 0.437399 (Rank: 7/10)

Analysis: Moderate enthusiasm, reflecting a decent level of engagement and energy. **Speech Speed:**

Speech Speed

1. **Speech Speed:** -0.888693 (Rank: 6/10)

Analysis: Slow speech speed, indicating a deliberate speaking pace.
Dominant Emotions:

Emotional State

1. **Top 1 Emotion:** Neutral
2. **Top 2 Emotion:** Fear

Analysis: Predominantly neutral with some fear, reflecting a calm demeanor with occasional apprehension. **Comparative Analysis:**

- **Strengths:**
  - **Sentiment:** 5th in positive score, showing a generally positive outlook.

  - **Enthusiasm:** 7th highest, indicating a decent level of engagement.

- **Weaknesses:**
  - **Confidence Score:** 8th lowest, suggesting moderate self-assurance.
  - **Hesitant Score:** 4th highest, reflecting notable hesitation.
  - **Speech Speed:** 6th slowest, indicating a slower delivery pace.
- **Areas of Expertise:**
  - **Sentiment:** Balanced positive outlook with a moderate approach.
  - **Enthusiasm:** Decent level of engagement and energy.

Overall: Student 8 has a balanced sentiment and moderate enthusiasm but should work on increasing confidence and managing hesitation to improve communication effectiveness.

Student 9

Transcript Sentiment:

Sentiment Breakdown

1. **Positive Score:** 0.617353
2. **Negative Score:** 0.223949
3. **Neutral Score:** 0.158699

Area of Expertise based on the wordblob image:

- **Artificial Intelligence and Agritech:** The prominent keywords "artificial intelligence" and "Agritech" suggest a focus on using AI in agricultural technology.
- **Facial Expressions and Engineering:** The terms indicate an interest in applying AI for facial expression recognition, possibly related to human-computer interaction or surveillance.
- **Development and Society:** The emphasis on "development" and "society" implies a goal of using technology for societal impact, likely in rural or agricultural settings.

Confidence and Hesitation:

Confidence and Hesitation Scores

1. **Confidence Score:** 0.591842
2. **Hesitation Score:** 0.538732

Analysis: The student demonstrates a fair level of confidence but a relatively high hesitation score, indicating some uncertainty or nervousness during their speech. **Enthusiasm:**

Enthusiastic Score

1. **Enthusiastic Score:** 0.505152

Analysis: The student shows moderate enthusiasm, reflecting a positive and engaging tone during their speech. **Speech Speed:**

Speech Speed

1. **Speech Speed:** 1.047063

Analysis: The student's speech speed is the fastest among all, which might be due to their high hesitation score, possibly rushing through points due to nervousness or time pressure. **Dominant Emotions:**

Emotional State

1. **Top 1 Emotion:** Neutral
2. **Top 2 Emotion:** Happy

Analysis: The student's emotional state is predominantly neutral with an undercurrent of happiness, indicating a generally balanced but somewhat positive demeanor. **Comparative Analysis:**

- **Strengths:**
  - **Positive Score:** 8th highest, showing a relatively positive outlook.
  - **Enthusiasm Score:** 6th highest, reflecting good engagement.
- **Weaknesses:**
  - **Hesitation Score:** 8th highest, indicating notable hesitation.
  - **Speech Speed:** Fastest among all, which may impact clarity.
- **Areas of Expertise:**
  - **Enthusiasm:** Moderate enthusiasm shows a positive tone in communication.
  - **Speed Management:** Needs to work on pacing to improve clarity.

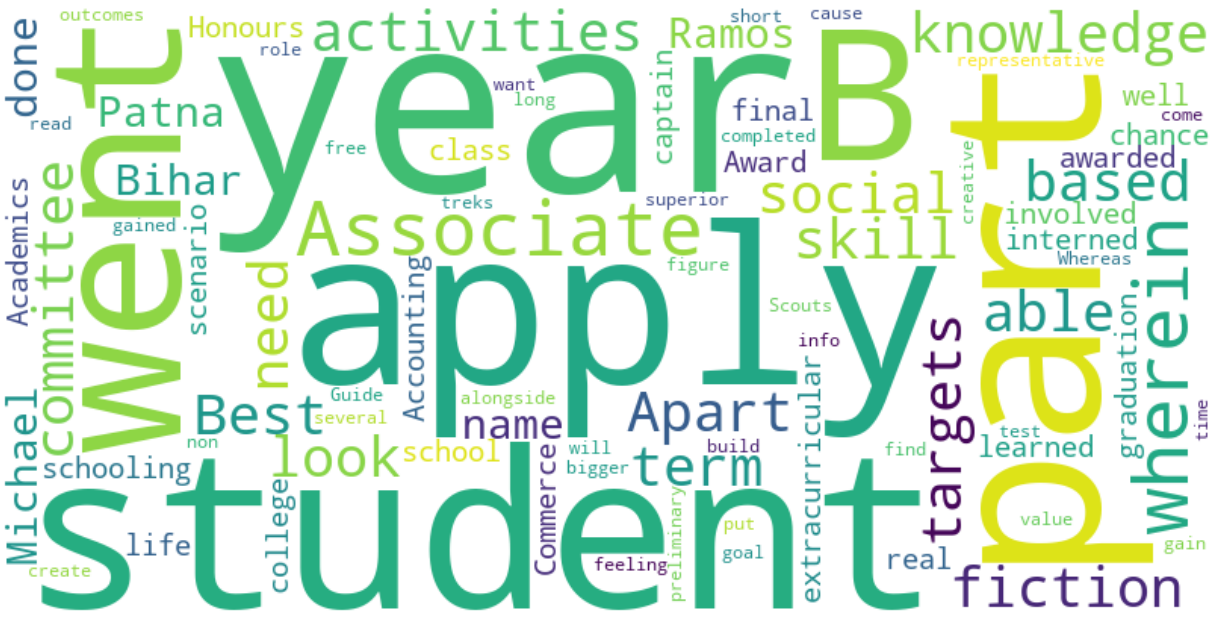
Overall: Student 9 displays a positive outlook and good enthusiasm but needs to address their hesitation and manage their fast speech pace to enhance overall communication effectiveness.

Student 10

Transcript Sentiment:

Sentiment Breakdown

1. **Positive Score:** 0.589267
2. **Negative Score:** 0.220948
3. **Neutral Score:** 0.189785



- **Student Life and Extracurriculars:** Keywords like "student," "apply," and "activities" indicate a focus on student engagement, academic involvement, and participation in extracurricular activities.
- **Commerce and Social Skills:** Terms like "Commerce," "skill," and "knowledge" suggest a background in commerce with an emphasis on developing social and professional skills.
- **Committee and Leadership:** The mention of "Committee" and "captain" indicates involvement in leadership roles, possibly within student organizations or extracurricular committees.

Confidence and Hesitation:

Confidence and Hesitation Scores

1. **Confidence Score:** 0.619852
2. **Hesitation Score:** 0.520637

Analysis: The student shows a good level of confidence, but the relatively high hesitation score suggests some nervousness or uncertainty in parts of their speech. **Enthusiasm:**

Enthusiastic Score

1. **Enthusiastic Score:** 0.325507

Analysis: The enthusiasm score is moderate, indicating a somewhat positive and energetic delivery, but lacking in high energy. **Speech Speed:**

Speech Speed

1. **Speech Speed:** 0.809499

Analysis: The student speaks at a fast pace, which could be related to their higher hesitation score, possibly due to nervousness or pressure. **Dominant Emotions:**

Emotional State

1. **Top 1 Emotion:** Sad
2. **Top 2 Emotion:** Fear

Analysis: The dominant emotions of sadness and fear suggest that the student may be experiencing emotional challenges or anxiety during their speech. **Comparative Analysis:**

- **Strengths:**
  - **Positive Score:** 9th highest, reflecting a generally positive

outlook.

  - **Confidence Score:** 4th highest, indicating strong self-assurance.

- **Weaknesses:**
  - **Enthusiasm Score:** 10th lowest, showing the least engagement.
  - **Speech Speed:** 9th fastest, which may affect clarity.
  - **Emotional State:** Dominant sadness and fear indicating underlying emotional challenges.
- **Areas of Expertise:**
  - **Confidence:** High self-assurance in communication.
  - **Speed Management:** Effective pacing needed to improve clarity and reduce nervousness.

Overall: Student 10 shows good confidence and a generally positive outlook but should focus on increasing enthusiasm, managing their speech speed, and addressing emotional challenges to improve communication effectiveness.

Conclusion

This comprehensive analysis of the video text dataset has revealed several significant patterns and relationships among emotional and communication attributes of students:

1. Strong correlation between confidence and positive emotions
2. Statistically significant difference between confident and hesitant scores
3. No significant relationship between positive emotions and speech speed
4. Hesitance and enthusiasm as the most important predictors of speech speed

These insights provide valuable information for understanding student behavior and can be utilized to optimize educational strategies, particularly in the context of video-based learning and assessment.

Future Work

To further enhance this analysis, consider the following directions for future research:

1. Conduct longitudinal studies to track changes in student communication patterns over time
2. Develop machine learning models to predict student performance based on communication attributes
3. Investigate the impact of different teaching methodologies on student communication styles
4. Expand the dataset to include a more diverse range of students and educational contexts

By pursuing these avenues, researchers and educators can gain even deeper insights into student behavior and develop more effective, personalized learning strategies.