

**Selected ERT:**

**Name:** Google Cloud Natural Language API (v2 version)

**Source:** <https://cloud.google.com/natural-language>

**Justification:** The selection of Google Natural Language API is based on its advanced capabilities in sentiment and emotion detection, which are crucial for analyzing emotional responses generated through text. This tool provides a reliable method for evaluating the emotional tone of interactions, making it suitable for this assignment.

**LLMA Used (ChatGPT):**

**Name:** ChatGPT (OpenAI, GPT-4o)

**Source:** <https://openai.com/chatgpt>

**Justification:** ChatGPT has been chosen for its flexibility and proficiency in handling a wide range of tasks. It is capable of performing both academic and non-academic functions, such as solving problems, answering questions, and engaging in informal conversation, which makes it highly suited for this assignment.

**Purpose and Context of Use:**

**Academic Activity:** Tool to solve complex mathematical or logical problems, offer in-depth analyses of scientific topics, and assist in the understanding of specific academic subjects. It will also be employed to simulate educational discussions and act as a tutor in areas such as literature, history, or technology.

Example: solve the quadratic equation

**Non-Academic Activity:** Tool for casual conversation, entertainment, and personal assistance. It aims to provide recommendations for everyday tasks (e.g., movies, books, restaurants), engage in informal chats, and assist with non-educational problem-solving (e.g., personal advice, hobby-related discussions). The goal is to provide a personalized and human-like interaction experience to the user, offering a balance between fun and utility.

Example: help you plan the vacation

**Data with Timestamps:**

**Think-Aloud Data:** A record of the thought processes articulated during the interaction. This data includes the user's verbalized thoughts or mental processes expressed during the interaction with the LLMA. As the user interacts, they may provide insight into their reasoning, questions, or clarifications. For example, when a user is solving a complex problem or trying to understand a scientific concept, their thought process may be articulated as they make decisions or formulate queries. Each statement or insight provided by the user is logged with a timestamp. The data will be collected using the Google API in csv or json format

## Academic

Timestamp	Thoughts (Transcription)
00:01	"I need to find the roots of this quadratic equation. Let's start with the formula."
00:15	"Okay, I'm substituting the values for a, b, and c into the quadratic formula."
00:30	"It looks like the discriminant is positive, so I'll have two real roots."
00:45	"I'm calculating the square root of the discriminant to find the next step."
01:00	"Let's divide the terms by 2a to isolate the solutions for x."
01:15	"The roots are 3 and -2. I'll double-check by plugging them back into the equation."
01:30	"Alright, the solutions seem correct. Now I'll move to analyze the graph of the quadratic function."
01:45	"I need to find the vertex of the parabola. Using the vertex formula, let's calculate $-b/(2a)$ ."
02:00	"The vertex is at (1, -4). Now, let's determine the axis of symmetry."
02:15	"The axis of symmetry is $x = 1$ . I'll also determine whether the parabola opens upwards or downwards."
02:30	"Since a is positive, the parabola opens upwards. This means the vertex is a minimum point."
02:45	"To find additional points on the graph, I'll plug $x = 0$ and $x = 2$ into the equation to calculate the y-values."
03:00	"I've calculated all the necessary points. Now I'll draw the graph to confirm the quadratic behavior visually."
03:15	"I should verify the symmetry of the graph by comparing points equidistant from the axis of symmetry."
03:30	"The graph looks correct. I'll summarize the key properties: roots, vertex, and symmetry."

## Non-academic

Timestamp	Thoughts (Transcription)
00:01	"I want to plan a relaxing beach vacation, maybe in January."
00:10	"It would be great if I could find somewhere warm, like Hawaii or the Caribbean."
00:25	"I wonder if there are good flight deals right now."
00:40	"I'll check for direct flights first to save time. Layovers can be so tiring."
00:55	"What about accommodation? I'd prefer something close to the beach, maybe an all-inclusive resort."
01:10	"I need to decide how long I want to stay. A week seems ideal, but maybe 10 days for extra relaxation."
01:25	"Let's check the weather in January for the top destinations on my list."
01:40	"I should compare prices between booking directly and using travel websites like Expedia or Kayak."
01:55	"It's also worth looking into any travel restrictions or requirements for international trips."
02:10	"I wonder if there are any festivals or events during that time. That could make the trip more interesting."
02:25	"What about activities? Snorkeling, hiking, or maybe just relaxing at the spa sounds appealing."
02:40	"Should I rent a car? It would give me more freedom to explore beyond the resort area."
02:55	"I'll need to budget for meals, activities, and souvenirs. Let's start estimating the total cost of the trip."
03:10	"Let me save some of these options in case I need to compare them later."
03:25	"Once I finalize the destination, I'll make a checklist for packing and trip preparations."
03:40	"I might also look up travel insurance to cover any unexpected changes or emergencies."
03:55	"I'll book the flights and hotel today if I find a good deal, just to

	lock in the prices."
--	----------------------

**Action Data:** A log of the actions performed, including queries entered and responses received. This data consists of the user's actions during the interaction, such as the queries typed into the chat, responses received from the LLMA, and subsequent follow-up actions. Every user action (typing a query or receiving a response) is logged with a timestamp.

#### Academic

Timestamp	Action	Response from LLMA
00:05	"Solve $x^2 + 3x - 4 = 0$ "	"The roots are $x = -4$ and $x = 1$ ."
00:20	"Explain how you derived the solution."	"I used the quadratic formula: $x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$ ."
00:35	"What is the discriminant in this case?"	"The discriminant is $b^2 - 4ac = 3^2 - 4(1)(-4) = 25$ ."
00:50	"Can you verify the solutions by substitution?"	"Yes. Substituting $x = -4$ gives $16 - 12 - 4 = 0$ . Substituting $x = 1$ gives $1 + 3 - 4 = 0$ ."
01:10	"What does the graph of this equation look like?"	"It's a parabola that opens upwards with roots at $x = -4$ and $x = 1$ , and a vertex at $(-1.5, -6.25)$ ."
01:25	"Find the vertex of the parabola."	"The vertex is at $x = -b/(2a) = -3/2 = -1.5$ . Substituting $x = -1.5$ into the equation gives $y = -6.25$ ."
01:45	"What is the axis of symmetry?"	"The axis of symmetry is $x = -1.5$ , the x-coordinate of the vertex."
02:00	"What is the y-intercept?"	"The y-intercept is at $x = 0$ . Substituting $x = 0$ into the equation gives $y = -4$ ."
02:20	"How does changing the value of 'a' affect the graph?"	"If ' $a$ ' > 0, the parabola opens upwards. If ' $a$ ' < 0, it opens downwards. Larger
02:40	"Explain why the equation has two roots."	"The discriminant is positive (25), which indicates two distinct real roots."

03:00	"What are some real-world applications of this equation?"	"Quadratic equations model projectile motion, area optimization, and parabolic reflectors."
03:20	"Can you solve $x^2 + 2x + 1 = 0$ ?"	"This is a perfect square trinomial. The root is $x = -1$ (repeated root)."
03:35	"Explain the perfect square trinomial case."	"When $b^2 = 4ac$ , the discriminant is 0, so the equation has a repeated root, $x = -b/2a$ ."
03:50	"How would the graph of $x^2 + 2x + 1$ look?"	"The parabola touches the x-axis at $x = -1$ (the vertex) and opens upwards."
04:10	"Summarize the properties of quadratic equations."	"Key properties: roots, vertex, axis of symmetry, y-intercept, and the effect of 'a' on the graph."

#### Non-academic

Timestamp	Action	Response from LLMA
00:05	"Suggest a warm destination for January."	"Hawaii, Maldives, and Cancun are great options."
00:15	"Can you help me compare Hawaii and Maldives?"	"Hawaii has more activities, while Maldives is more serene."
00:30	"Find me a flight to Hawaii."	"There's a flight for \$500 on January 10th from LAX."
00:45	"What are the best beaches in Hawaii?"	"Some top beaches include Waikiki Beach, Lanikai Beach, and Hanauma Bay."
01:00	"What's the weather like in Hawaii in January?"	"Hawaii in January is warm, with temperatures ranging from 75°F to 80°F, and occasional light rain."
01:15	"Are there any festivals in Hawaii in January?"	"The Maui Whale Festival and the Pacific Islands Arts Festival are held in January."
01:30	"Recommend activities for a	"Family-friendly activities

	family in Hawaii."	include snorkeling at Hanauma Bay, visiting the Dole Plantation, and a luau show."
01:45	"Can you help me find an all-inclusive resort?"	"The Four Seasons Resort in Maui and Hilton Hawaiian Village in Waikiki are excellent options."
02:00	"How much should I budget for 5 days in Hawaii?"	"For 5 days, a budget of \$3,000–\$5,000 is reasonable, including flights, accommodations, and activities."
02:20	"What documents do I need for the trip?"	"For U.S. citizens, no passport is needed for Hawaii. Just ensure you have a valid ID for travel."
02:40	"Suggest restaurants in Honolulu."	"Top-rated restaurants include Alan Wong's, Helena's Hawaiian Food, and Duke's Waikiki."
03:00	"How can I save money while visiting Hawaii?"	"Consider booking off-season, choosing vacation rentals, and using public transport or shared tours."
03:20	"What are the best spots for snorkeling?"	"Hanauma Bay, Molokini Crater, and Shark's Cove are excellent snorkeling locations in Hawaii."
03:40	"What souvenirs should I buy in Hawaii?"	"Popular souvenirs include macadamia nuts, Hawaiian shirts, ukuleles, and Kona coffee."
04:00	"Can you help me plan a day-by-day itinerary?"	"Day 1: Relax on Waikiki Beach. Day 2: Explore Diamond Head. Day 3: Snorkel at Hanauma Bay. Day 4: Visit Pearl Harbor. Day 5: Luau experience."

**Emotion Data:** Sentiment analysis results from Google Natural Language API. This data is generated by the Google Cloud Natural Language API, which performs sentiment and emotion

analysis on the user's input. The tool detects emotions like joy, frustration, or confusion, and provides an overall emotional score for each interaction. Emotion data is recorded at key moments throughout the conversation, with a timestamp for each emotional response.

### Academic

Timestamp	Input Text	Emotion	Sentiment Score
00:10	"I feel like I understand this now."	Joy/Confidence	+0.85
00:35	"This part is confusing."	Frustration	-0.4
00:50	"I don't understand this step."	Confusion	-0.5
01:00	"Okay, now it makes sense."	Relief	+0.6
01:15	"I still feel unsure about this concept."	Doubt	-0.3
01:30	"I'm starting to get the hang of it."	Hopefulness	+0.65
01:45	"This is actually quite fun!"	Enjoyment	+0.9
02:00	"I feel like I'm close to solving this."	Confidence	+0.8
02:15	"I need a bit more clarification here."	Uncertainty	-0.2
02:30	"This is frustrating."	Frustration	-0.6
02:45	"I think I got it now."	Relief/Confidence	+0.85
03:00	"This part is tricky, I'm not sure."	Anxiety	-0.45
03:15	"I'll try again with a different approach."	Determination	+0.5
03:30	"I did it! This is great!"	Joy	+1.0

### Non-academic

Timestamp	Input Text	Emotion	Sentiment Score
-----------	------------	---------	-----------------

00:05	"Suggest a warm destination for January."	Curiosity	+0.4
00:15	"Can you help me compare Hawaii and Maldives?"	Neutral	+0.2
00:30	"Find me a flight to Hawaii."	Joy	+0.8
00:45	"I wonder if there are any good deals on hotels in Hawaii."	Curiosity	+0.5
01:00	"What is the best time of year to visit the Maldives?"	Interest	+0.6
01:20	"I'm getting excited about this trip!"	Excitement	+0.9
01:35	"Should I book a direct flight or a layover?"	Uncertainty	+0.1
01:50	"I think I'll go for the Maldives after all."	Decision-making	+0.7
02:10	"Let's check for activities to do in Hawaii."	Excitement	+0.8
02:25	"How do I get from the airport to the hotel?"	Mild Frustration	-0.2

### Emotional trajectory visualised.

**Data Collection:** Emotions are analyzed at different points during the conversation, capturing emotions like joy or frustration, with timestamps for each user input.

**Visualization:** X-axis represents time (timestamps for when each input was processed).

Y-axis represents emotional intensity, shown by the sentiment score returned from the sentiment analysis (a range from -1 for negative sentiment to +1 for positive sentiment).

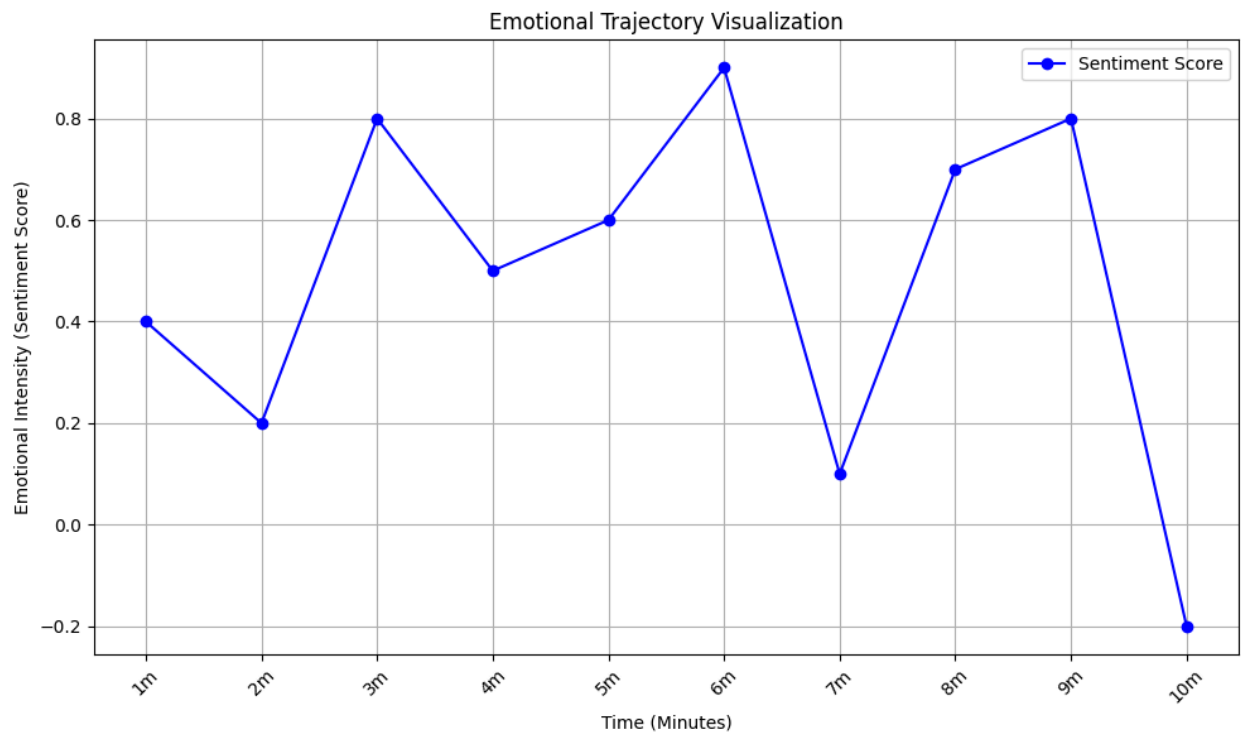
Line graph with sentiment scores plotted at each timestamp shows the emotional trajectory of the user's inputs over time.

**Analysis:** Peaks in positive emotions suggest good engagement, while drops indicate frustration or disengagement. This helps identify successful moments or areas for improvement in the LLMA's responses.



**Insights:** Emotional trends provide insights into user satisfaction and trust, showing when the system fails or succeeds in keeping the user engaged.

Sample of the graph



<b>LLMA (ChatGPT - OpenAI, GPT-4). Academic activity. Trust measures.</b>	Strongl y disagre e	Rather disagre e	Neither disagre e or agree	Rather agree	Strongl y agree	No respons e
<b>Accuracy and Reliability.</b> Does ChatGPT consistently provide correct and relevant responses to academic queries (e.g., solving complex problems, providing in-depth analyses)?	(1)	(2)	(3)	(4)	(5)	( )
<b>Transparency.</b> Are the limitations of ChatGPT clear when it comes to academic tasks (e.g., inability to access real-time data or the internet)? Can users understand how the model operates?	(1)	(2)	(3)	(4)	(5)	( )
<b>Predictability.</b> Is the behavior	(1)	(2)	(3)	(4)	(5)	( )

of ChatGPT predictable when responding to academic queries? Does the system provide consistent and relevant information in academic contexts?						
<b>Security and Privacy.</b> Does ChatGPT handle sensitive academic data securely? Are there clear privacy policies in place regarding data usage in academic interactions?	(1)	(2)	(3)	(4)	(5)	( )
<b>Perceived Expertise.</b> Does the user trust ChatGPT based on its displayed academic knowledge and problem-solving ability? Is the user confident in the model's ability to explain complex concepts or provide reliable answers?	(1)	(2)	(3)	(4)	(5)	( )

<b>LLMA (ChatGPT - OpenAI, GPT-4). Non-academic activity. Trust measures.</b>	Strongly disagree	Rather disagree	Neither disagree or agree	Rather agree	Strongly agree	No response
<b>Accuracy and Reliability.</b> Does ChatGPT consistently provide accurate and relevant answers to non-academic queries (e.g., personal advice, recommendations)?	(1)	(2)	(3)	(4)	(5)	( )
<b>Transparency.</b> Are the limitations of ChatGPT clear when used for non-academic tasks (e.g., entertainment, personal assistance)? Can users understand how the model works?	(1)	(2)	(3)	(4)	(5)	( )
<b>Predictability.</b> Is the behavior of ChatGPT predictable in casual or non-academic interactions? Does the system	(1)	(2)	(3)	(4)	(5)	( )

provide consistent and appropriate responses?						
<b>Security and Privacy.</b> Does ChatGPT ensure that users' personal information is securely handled during non-academic interactions? Are privacy policies in place for data usage?	(1)	(2)	<b>(3)</b>	(4)	(5)	( )
<b>Perceived Expertise.</b> Does the user trust ChatGPT based on its proficiency in providing personal advice, recommendations, or engaging in casual conversation?	(1)	(2)	(3)	(4)	<b>(5)</b>	( )

<b>ERT (Google Cloud Natural Language API - v2 version). Trust Measures</b>	Strongly disagree	Rather disagree	Neither disagree or agree	Rather agree	Strongly agree	No response
<b>Accuracy of Emotion Detection.</b> Does the ERT reliably detect emotions from user input in both academic and non-academic contexts? How accurate is the emotion score?	(1)	(2)	(3)	(4)	<b>(5)</b>	( )
<b>Relevance.</b> Are the emotional responses generated by the ERT aligned with the user's context, whether academic or non-academic?	(1)	(2)	(3)	<b>(4)</b>	(5)	( )
<b>Transparency.</b> Can users understand how the Google Natural Language API performs its emotion analysis? Are the methods and processes clearly communicated to users?	(1)	(2)	(3)	(4)	<b>(5)</b>	( )
<b>Consistency.</b> Does the ERT provide consistent emotion analysis over time, regardless	(1)	(2)	(3)	<b>(4)</b>	(5)	( )

of the interaction type (academic or non-academic)? Are emotional responses steady and reliable?						
<b>Impact on Interaction.</b> Does the emotional detection feature improve user engagement and trust in both academic and non-academic interactions? Does it enhance the overall experience with the LLMA?	(1)	(2)	(3)	<b>(4)</b>	(5)	( )

Hoffman et al's (2023) "the Trust Scale for the AI Context" (TAI)

Each item is rated by a 5-point Likert scale: 1: I agree strongly, 2: I agree somewhat, 3: I'm neutral to

it, 4: I disagree somewhat, 5: I disagree strongly.