香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen

Large Language Models for Emotion Evolution Prediction

Clement Leung
Zhifei Xu

香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen

Part 1 – Introduction

Part 2 – Background & Motivation

Part 3 – Methodology

Part 4 – Experiments & Results

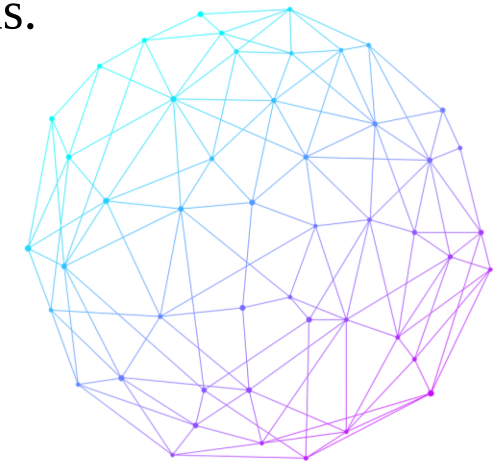
Part 5 – Conclusion & Discussion

Introduction

- This study is to investigate how emotions can be recognized, classified, and predicted using multimodal data.

- We categorize emotional states into positive (competent) and negative (incompetent) groups to assess emotional competency and developed a stochastic model to dissect the dynamics and progression of emotional states, offering a framework to study how emotions evolve and their sustained impact on individuals.

- We also explore the potential of Large Language Models to assist and facilitate accurate emotion recognition in zero-shot situations.



香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen

Background & Motivation

- Emotions play a critical role in human interactions and decision-making, especially in safety-sensitive roles such as pilots, surgeons, and emergency responders. Accurate emotion recognition can help in monitoring and improving performance, ensuring safety, and enhancing overall well-being.

- For example: a pilot who attempted to shut down the engines of an airplane mid-fligh

Background & Motivation

What is emotion recognition?

- Emotion recognition refers to the technological process of identifying and interpreting human emotions.
- People vary significantly in their ability to accurately recognize the emotions of others, therefore, it leads to emotion recognition field.
- Influences individual preferences, decisions, and behaviors.

Emotion recognition and prediction technology has garnered significant interest due to its potential to enhance safety, support mental health, and improve user experiences. This field has been acknowledged as a critical factor in enhancing human safety and has thus been the subject of extensive research.

Background & Motivation

◆ Emotion recognition and prediction techniques have developed into two main types:

Single-modal system - It utilize a single type of data, such as text, speech, or facial expressions, to detect and predict emotions.

Multi-modal system - MER systems enhance accuracy and robustness by integrating data from multiple sources, such as facial expressions, speech patterns, and physiological signals.

Part 1 – Introduction

Part 2 – Background & Motivation

Part 3 – Methodology

Part 4 – Experiments & Results

Part 5 – Conclusion & Discussion

Methodology

- Emotions are dynamic and continuously transition from one state to another
  - Accurately predicting the emotional states of individuals is essential, particularly when assigning work rosters or coordinating tasks that require high concentration and emotional stability.
  - Such emotional assessments are key to preventing potential accidents or errors that could arise from impaired judgment caused by adverse emotional states. For instance, a worker dealing with unmanaged anger or extreme sadness might not have the necessary focus or decision-making capacity to safely operate machinery or make critical, split-second decisions.

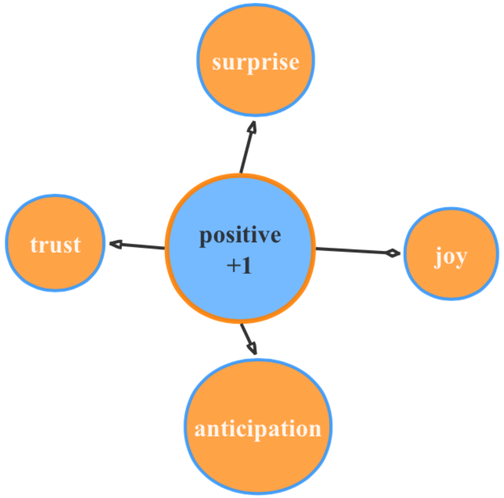
Methodology

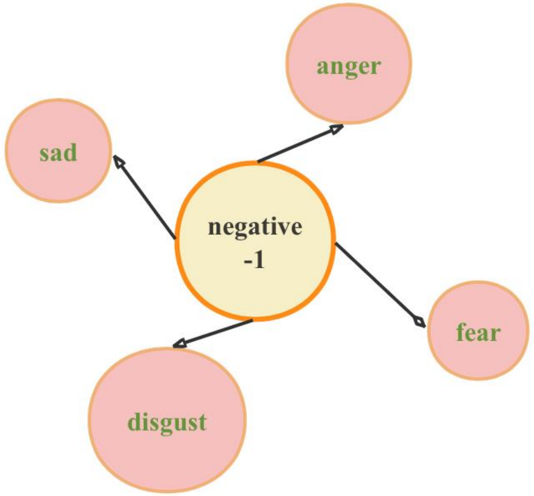
- We apply Ekman's model, which identifies six primary emotions.

- Positive emotion(+1): happiness, surprise, neutral

- Negative emotion(-1): sadness, disgust, anger

- Then, we propose a key element in this model; the emotion stability factor, λ





Methodology

- We represent the emotional state at time **t** by **S(t)**; **t** is time, and **S(t)** is the person's emotional change with time. As indicated, **S(t)** can be take on the values:

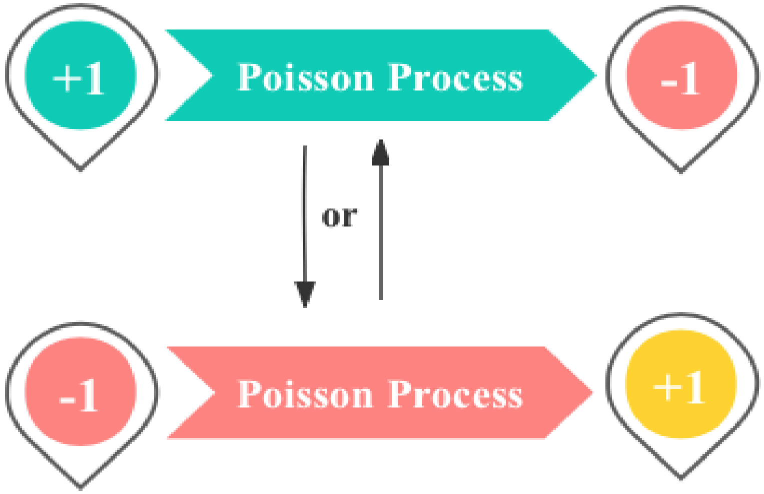
$$S(t) = +1 \qquad\qquad S(t) = -1$$

- First, let S(0)=1 then, we represent the transition time points by a Poisson Process. Now, S(t)=1 if the number of transitions in the time interval (0,t) is even, and S(t)=-1 if this number is odd. Therefore,

$$P[S(t)=1|S(t)=0] = p_0 + p_2 + p_4 + ... + ...,$$

where p_k is the number of Poisson points in (0,t) with parameter

$$p_k = e^{-\lambda t}\frac{(\lambda t)^k}{k!}$$



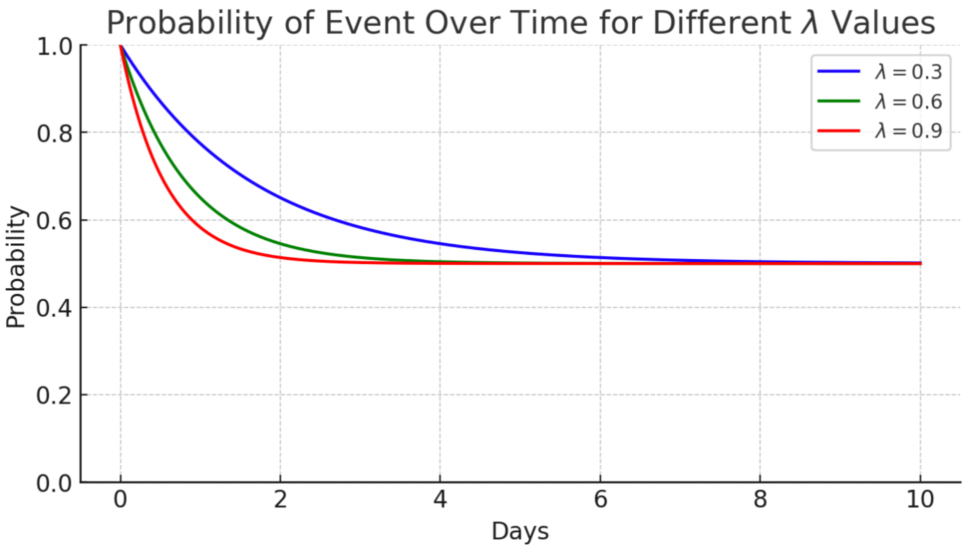
That is,

$$P[S(t)=1|S(0)=0] = e^{-\lambda t}[1 + \frac{(\lambda t)^2}{2!} + \frac{(\lambda t)^4}{4!} ... + ...]$$
$$= e^{-\lambda t}\cosh \lambda t$$

with

$$\cosh(t) = \frac{e^{-t}+e^{t}}{2}$$

Methodology



$$e^{-\lambda t}\cosh \lambda t$$

According to this image, we can understand the changes of emotional states of different individuals under different conditions.

Analyzing the day-to-day probabilities:

first day -- over 50\% probability of positive emotions -- probabilities ranging from 60% to 80%

second day -- probabilities around the 50%

fourth day -- natural equilibrium state in emotional dynamics

Methodology

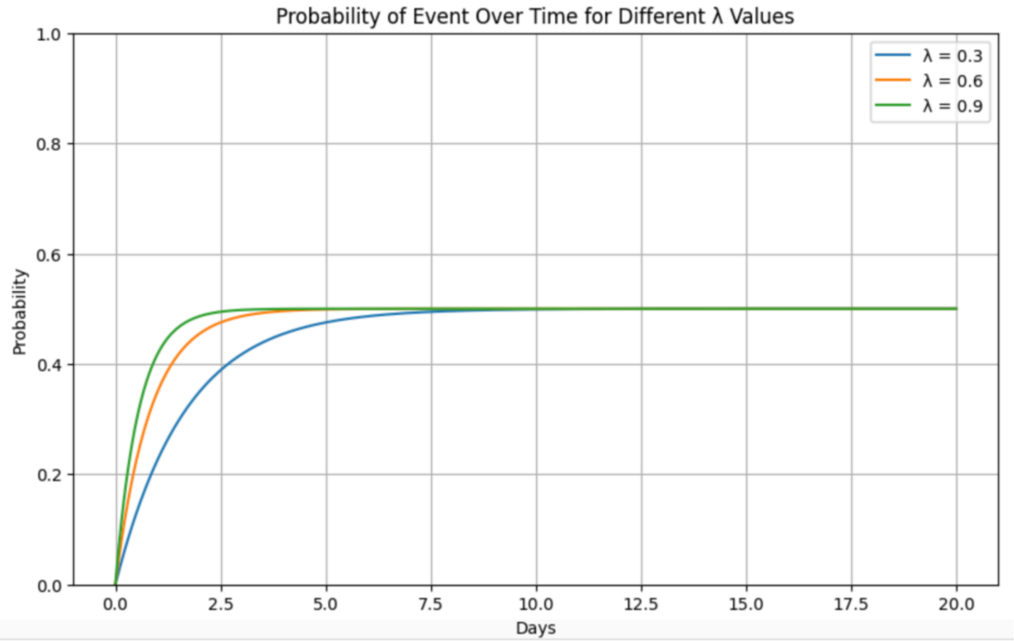
- Now, S(t)=-1 if the number of points in the time interval (0,t) is odd; we have:

$$P[S(t) = -1|S(0) = 0]] = e^{-\lambda t}[1 + \frac{(\lambda t)^3}{3!} + \frac{(\lambda t)^5}{5!} ... + ...]$$
$$= e^{-\lambda t} \sinh \lambda t$$

- This series is equivalent to the series expansion of sinh(t):

$$\sinh(t) = \sum_{k=0}^{\infty} \frac{t^{2k+1}}{(2k+1)!}$$

Methodology



$$e^{-\lambda t}\sinh \lambda t$$

- Provides the dynamics of probability over time.
- It is clear from the figure that emotional stability factors with different attenuation constants can significantly affect the duration and intensity of probabilistic states over several days.

- λ =0.3 The probability starts out strong but decays slowly, indicating a lingering effect or a state of slow weakening.
- λ =0.6 The probability are dropping fast. This indicates that the state dissipates faster and the emotional intensity or probability of staying in the same state decreases faster.
- λ =0.9 The probability goes down even faster. The curve drops sharply, reflecting situations in which a negative emotional state or other similar situation dissipates very quickly.

Methodology

- State Space: 8 emotions (Ekman + expansions like trust, anticipation)
- Markov Transition Matrix P

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{18} \\ \vdots & \vdots & \ddots & \vdots \\ p_{81} & p_{82} & \cdots & p_{88} \end{bmatrix}$$

- Smaller λ_i $\Rightarrow$ more stable (e.g., positive emotions)
- Larger λ_i $\Rightarrow$ less stable (e.g., negative emotions)

香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen

Part 1 – Introduction

Part 2 – Background & Motivation

Part 3 – Methodology

Part 4 – Experiments & Results

Part 5 – Conclusion & Discussion

Data & Experiment

- **Dataset**
  - Emotion Detection - This dataset contains 35,685 grayscale images, each with a resolution of 48x48 pixels.
  - Facial Expressions Training - The AffectNet database hosts a substantial collection of facial images, and the image has been reduced to 96x96 pixels.
  - Natural Human Face Images for Emotion Recognition - Each image is formatted as a 224x224 pixel grayscale image. These images are sourced from publicly accessible online platforms including Google, Unsplash, and Flickr.
  - MELD: Multimodal EmotionLines Dataset it incorporates over 1,400 dialogue sequences and encompasses approximately 13,000 spoken exchanges, all extracted from the popular TV series "Friends."

Data & Experiment

◆ Tasks Defined

- Image + scenario prompt → measure ChatGPT4's predicted emotional outcome. This task evaluates how well ChatGPT4 can predict emotional changes based on given scenarios.
-
- Image + different emotional sentences → measure ChatGPT4's classification. This task assesses ChatGPT4's ability to classify emotions based on images and associated text.

Experiments & Results

- Use TPR, FPR and ROC to evaluation Large Language Models. The model's performance varies with different threshold settings, emphasizing the need for optimal threshold selection.

- Negative emotions: 79.4% accuracy in negative contexts, 72.8% in positive contexts.

- While effective in identifying surprise, Large Language Models struggles to determine its sentiment, often defaulting to a neutral classification.

Experiments & Results

- ChatGPT4 generally achieves higher accuracy for positive emotions compared to negative emotions. The overall accuracy ranges from 70% to 95%, depending on the context and specific emotions. Common confusion occurs in distinguishing “disgust” from “sadness” or “anger,” indicating a need for further refinement in handling negative emotions.

- ROC curves are used to visualize the model’s performance in distinguishing between positive and negative emotions. The AUC values indicate the model’s discriminative power, with higher values suggesting better performance. The results show strong sensitivity for happiness, neutral, and surprise, while specificity is lower for negative emotions unless more detailed information is provided.

- The model sometimes under- predicts transitions from negative to positive emotions. For example, when an initial expression is strongly negative, the model often “stays negative” even when a positive scenario is presented.
- By varying the (λ) values, the model can adjust how quickly an emotion can switch from negative to positive, providing a more nuanced understanding of emotional dynamics.

Experiments & Results

Emotion	Parameter	Positive Situation	Negative Situation
Anger	accuracy	68.30%	73.30%
	sensitivity	NaN	NaN
	specificity	68.30%	73.30%
Disgust	accuracy	78.30%	85.00%
	sensitivity	NaN	NaN
	specificity	78.30%	85.00%
Happiness	accuracy	91.70%	83.30%
	sensitivity	91.70%	83.30%
	specificity	NaN	NaN
Neutral	accuracy	86.70%	83.30%
	sensitivity	86.70%	83.30%
	specificity	NaN	NaN
Sad	accuracy	71.70%	80.00%
	sensitivity	NaN	NaN
	specificity	71.70%	80.00%
Surprise	accuracy	85.00%	90.00%
	sensitivity	85.00%	90.00%
	specificity	NaN	NaN
Negative	accuracy	72.80%	79.40%
	sensitivity	NaN	NaN
	specificity	72.80%	79.40%
Positive	accuracy	87.80%	85.60%
	sensitivity	87.80%	85.60%
	specificity	NaN	NaN

Result of Four Different Situations with Different Emotions

Emotion	Anger sentence	disgust Sentence	Happiness sentence	Neutral Sentence	Sad sentence	Surprise sentence
Anger	70.00%	86.70%	86.70%	86.70%	86.70%	83.30%
Disgust	60.00%	70.00%	60.00%	56.70%	83.30%	56.70%
Happiness	70.00%	96.70%	1	96.70%	96.70%	96.70%
Neutral	76.70%	86.70%	96.70%	96.70%	90.00%	90.00%
Sad	63.30%	76.70%	76.70%	76.70%	86.70%	86.70%
Surprise	73.30%	86.70%	96.70%	96.70%	93.30%	96.70%

Result of Six Different Categories Emotional Sentences

香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen

Part 1 – Introduction

Part 2 – Background & Motivation

Part 3 – Methodology

Part 4 – Experiments & Results

Part 5 – Conclusion & Discussion

Conclusions

- Developed model with emotion stability factor λ to understand individual differences in emotional dynamics, enabling tailored interventions and treatments.
- Mathematical modeling can predict emotional shifts, allowing proactive strategies to improve mental health and wellbeing.
- Experimental results show that in positive emotions, the predictions of the large language model are consistent with the trend of our model for emotion prediction changes and they yield current accuracy in relation to explain in emotion state.

香港中文大學（深圳）
The Chinese University of Hong Kong, Shenzhen

Thanks !