AI Emotional Companion

Your Empathetic AI Friend, Always There
Anytime, without judgment

Intro to Deep Learning

Micah Baldonado, Prajwal Kumar, Aryaman Shandilya, Peter Ragone

(Slidesfree, 2018)

Agenda

01 Problem Statement

02 Solution and Overall Model

03 Speech Emotion Recognition Demo

04 Multimodal AI Companion Demo

2 demos!

The Problem

- 1 in 5 adults in the U.S. experiences mental illness each year (NAMI, 2023)
- Access to therapy is a challenge
  - Cost ($100–$200 per session) (Healthline)
  - Time (Often weeks)
  - Stigma

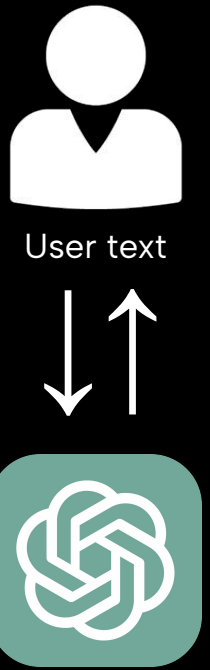
Solution

AI Emotional Companion

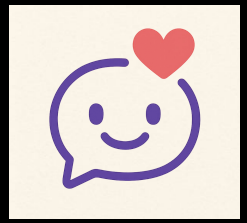
- Offers immediate emotional relief for those without easy access to a therapist

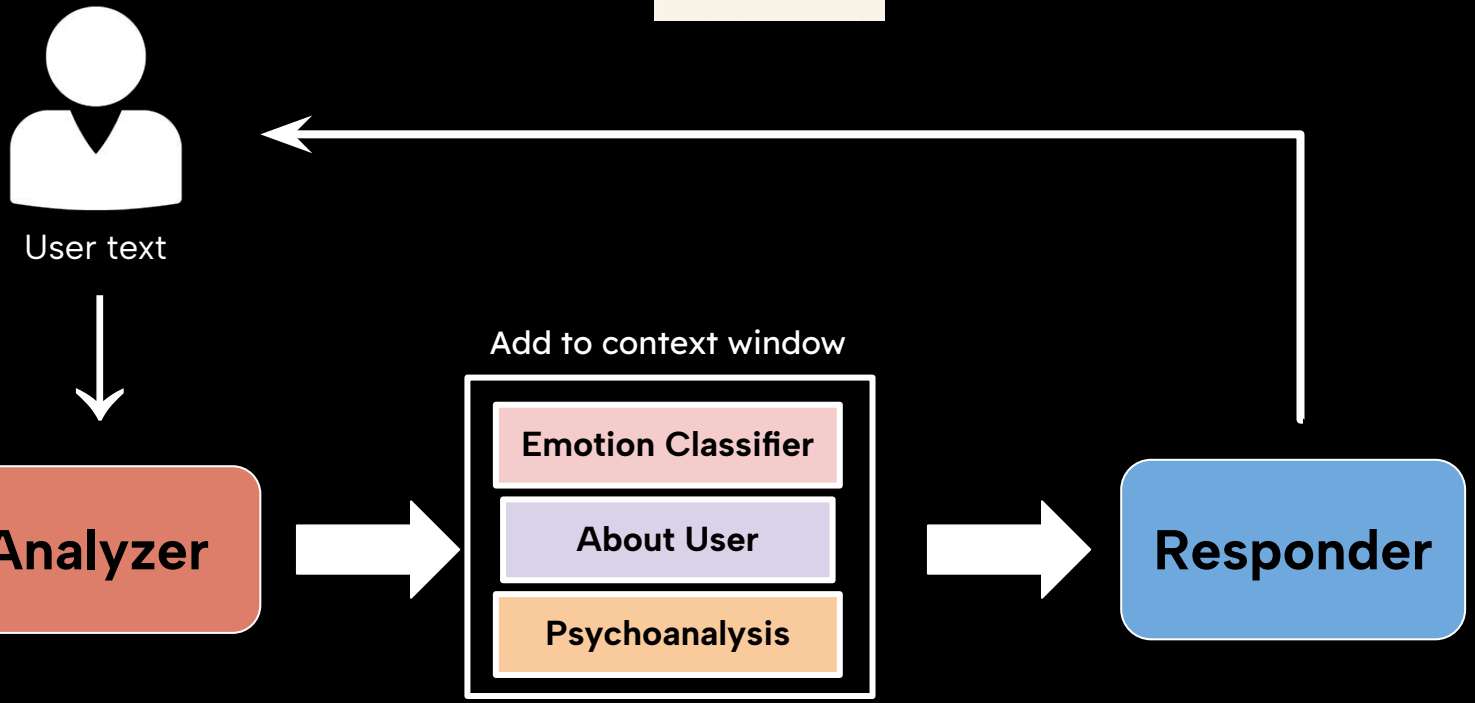
*Not a replacement for therapy, but a readily available tool for everyday emotional well-being

Normal Call to LLM:



Our New




User text
Analyzer
Add to context window
Emotion Classifier
About User
Psychoanalysis
Responder

Does the new framework increase an LLM's capacity for empathy?

P1 ChatGPT

P2 AI Emotional Companion

VS

Image source: Super Smash Bros. Ultimate, Nintendo

Putting Results in Context

- **Evaluated model responses on a HuggingFace dataset** (Reddit mental health posts)

Putting Results in Context

- **Evaluated model responses on a HuggingFace dataset** (Reddit mental health posts)
- Exact same prompt (the posts)

Putting Results in Context

- **Evaluated model responses on a HuggingFace dataset** (Reddit mental health posts)
- Exact same prompt (the posts)
- Exact same role (act as a therapist)

Putting Results in Context

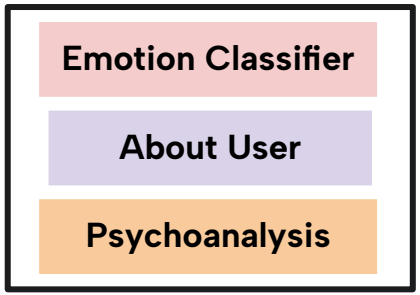
- **Evaluated model responses on a HuggingFace dataset** (Reddit mental health posts)
- Exact same prompt (the posts)
- Exact same role (act as a therapist)
- Zero-shot

Putting Results in Context

- **Evaluated model responses on a HuggingFace dataset** (Reddit mental health posts)
- Exact same prompt (the posts)
- Exact same role (act as a therapist)
- Zero-shot
- Only difference:
  - Context Window

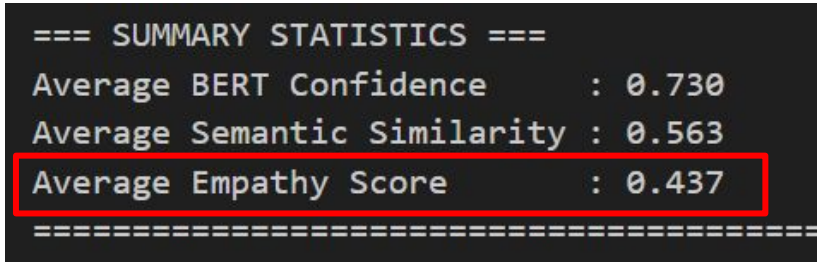
Putting Results in Context

- **Evaluated model responses on a HuggingFace dataset** (Reddit mental health posts)
- Exact same prompt (the posts)
- Exact same role (act as a therapist)
- Zero-shot
- Only difference:
  - Context Window



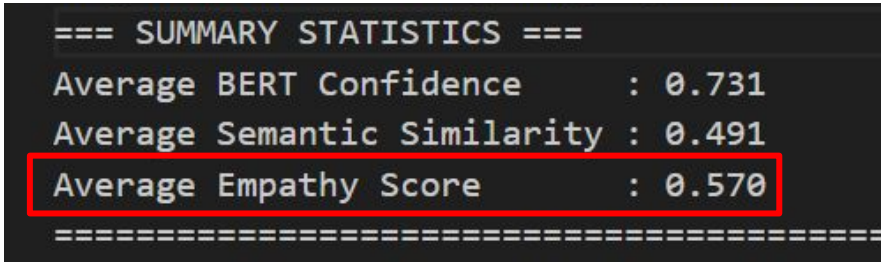
Results (zero-shot training with same input)

Vanilla ChatGPT



Empathy Score: Significantly Lower

Emotional Companion AI



Empathy Score: Significant improvement

While more rigorous testing is needed, our team is looking to publish our results. **Simply by seeing the same information from a different viewpoint, a LLM can respond more empathetically.**

Why not take this one step further?

Multimodal AI Emotional Companion

Old Framework :

User text

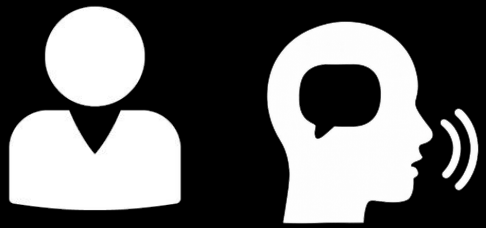
Analyzer

Add to context window

- Text Emotion
- About User
- Psychoanalysis

Responder

Multimodal :



User text

Add to context window

Speech Emotion

Text Emotion

About User

Psychoanalysis

Analyzer

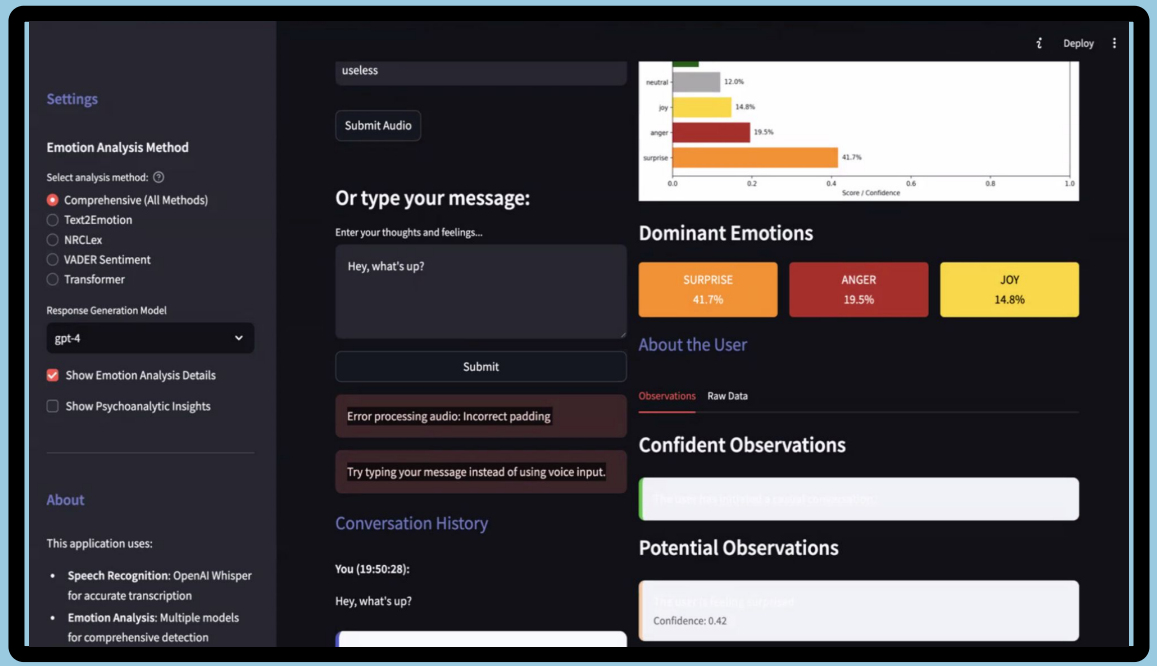
Responder

Speech Emotion Recognition Demo

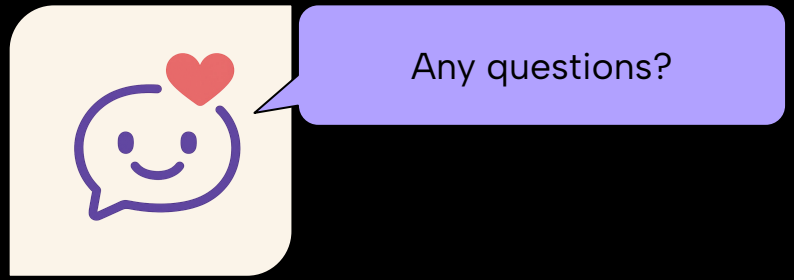
(Deep learning component of project)

Speech Emotion Recognition (SER) Demo

Multimodal AI Companion Demo


Deploy
Settings
Emotion Analysis Method
Select analysis method:
Comprehensive (All Methods)
Text2Emotion
NRCLex
VADER Sentiment
Transformer
Response Generation Model
gpt-4
Show Emotion Analysis Details
Show Psychoanalytic Insights
About
This application uses:
Speech Recognition: OpenAI Whisper for accurate transcription
Emotion Analysis: Multiple models for comprehensive detection
useless
Submit Audio
Or type your message:
Enter your thoughts and feelings...
Hey, what's up?
Submit
Error processing audio: Incorrect padding
Try typing your message instead of using voice input.
Conversation History
You (19:50:28):
Hey, what's up?
neutral
joy
anger
surprise
12.0%
14.8%
19.5%
41.7%
0.0
0.2
0.4
0.6
0.8
1.0
Score / Confidence
Dominant Emotions
SURPRISE
41.7%
ANGER
19.5%
JOY
14.8%
About the User
Observations
Raw Data
Confident Observations
Potential Observations
Confidence: 0.42

Thank you! :)



References

1. National Alliance on Mental Illness. (2023). Mental health by the numbers. Retrieved from https://www.nami.org/about-mental-illness/mental-health-by-the-numbers/#:~:text=22.8%25%20of%20U.S.%20adults%20experienced,represents%201%20in%205%20adults
2. Caplan, E. (2024, February 27). How much does therapy cost? A deep dive into prices. Healthline. https://www.healthline.com/health/how-much-does-therapy-cost-a-deep-dive-into-prices

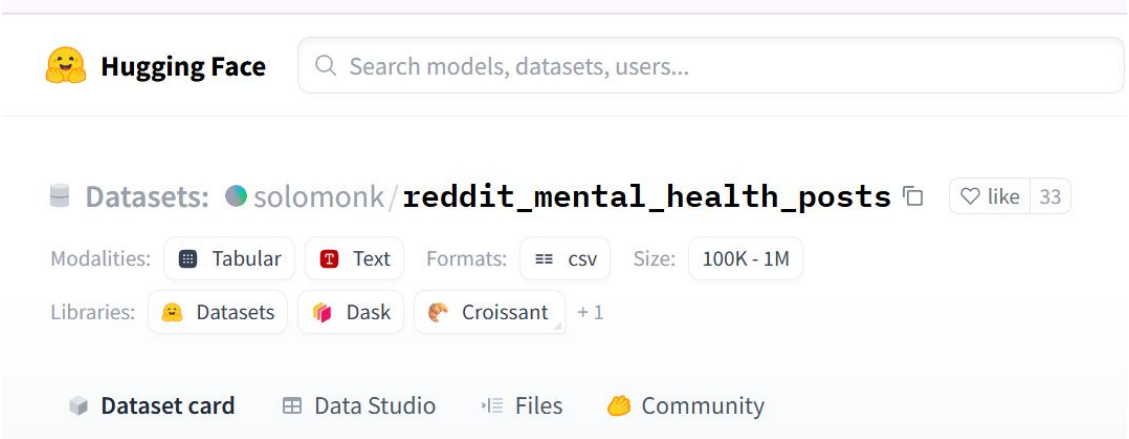
Image Citations

1. Slidesfree. (2018, October 10). Spectacular cloud PPT background images [Image]. Freeppt7.com. https://www.freeppt7.com/backgrounds/3845.html
2. OpenAI. (n.d.). ChatGPT logo [Logo]. https://openai.com
3. Nintendo. (2018). Super Smash Bros. Ultimate [Screenshot]. Nintendo.

Additional Slides

For reference during Q&A

Dataset: Hugging Face: Previous Reddit Posts



N = 44

- Scraped from posts by users on Reddit posting about emotional distress
- Used this to simulate real input queries for the model.

*Note: I used different dataset that has since been scraped from the internet (it still exists on hugging face. Previously available at https://huggingface.co/datasets/reddit/mental_health

Evaluation Metrics: How to model effectiveness?

- Need empathy from output based on input
- Unfortunately, most open-source empathy evaluators are response-focused only
    - Except notably EmpGPT-3, which I was unable to implement for this project.
- Solution: focus on multiple metrics that compliment each other

Evaluation Metrics

- Empathy Score
  - How empathetic-sounding the response is (self-explanatory)
- BERT Score
- Semantic Similarity

BERTScore vs. Semantic Similarity

Summary:

- **BERTScore** → fine-grained matching, like a *logical check*.
- **Semantic Similarity** → big-picture alignment, like a *topic check*.

(more complicated than this, sticking with simplicity)

Example:
Input: "I lost my job."
Output A: "That must be hard." → High BERTScore
Output B: "I like cake." → Low BERTScore

Example:
Input: "My dad just died."
Output A: "That must be hard." → High Semantic Similarity
Output B: "You should try skydiving." → Low Semantic Similarity

BERTScore vs. Semantic Similarity: Alternate Viewpoint

🤖 BERT is trained to:

> Understand the context of a word in a sentence by predicting masked words.

- It captures **contextual word meaning**, not just similarity between words.
- Example: The word "cold" in
  - "I have a cold" vs.
  - "It's cold outside"
    will have different embeddings.

Takeaway: BERT isn't directly trained to judge similarity between two sentences — it's trained to deeply **understand language structure**. But you can **use its embeddings** to **compare sentences** via cosine similarity — and this becomes a kind of semantic similarity.

🧠 Semantic Similarity models (like SBERT):

> Are fine-tuned versions of models like BERT to explicitly compare sentence meanings.

- They're trained on **sentence pairs** labeled as "similar" or "not similar".
- Goal: Learn embeddings such that semantically similar sentences are **close in vector space**.

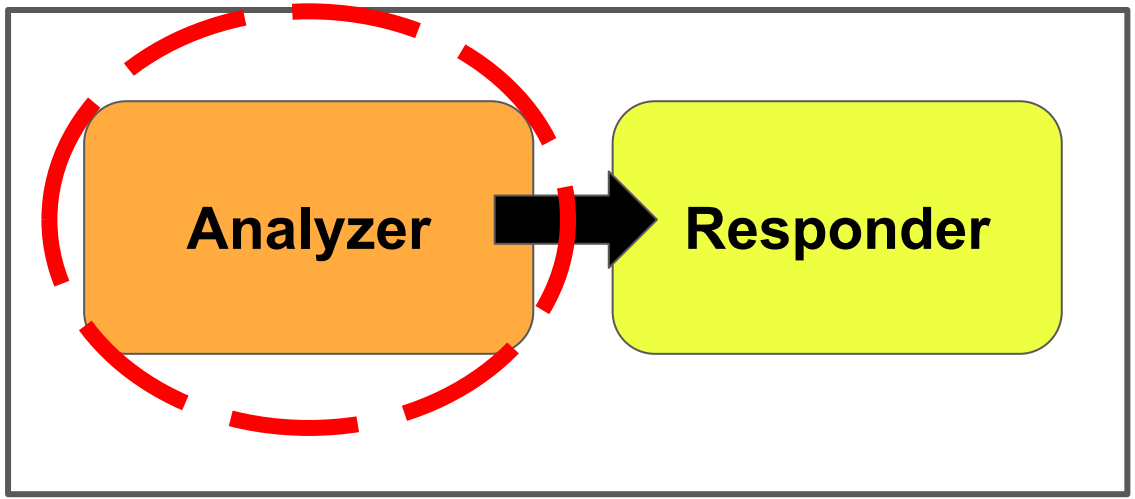
🧠 2. Why it acts like a logic check

BERTScore rewards:

- Lexical **relevance**: Are the response tokens *talking about the same things*?
- **Contextual fit**: Are the words used in similar situations?

Analyzer

- Emotion Classifier
- About User (Certain, Uncertain [short term prob, long term prob])
- Psychoanalysis



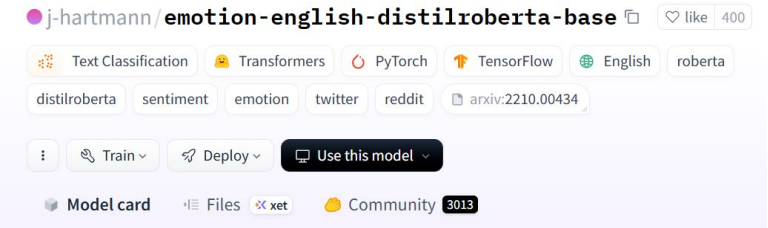
Analyzer: Emotion Sentiment Classifier

NRCLex (Top Emotions):
negative: 0.2631578947368421
sadness : 0.2631578947368421

VADER Sentiment (Polarity Scores):
neg : 0.20
neu : 0.65
pos : 0.15
compound: -0.52

Transformer-Based Emotion Classification (GoEmotions-like):
anger : 0.0062
disgust : 0.0779
fear : 0.0284
joy : 0.0032
neutral : 0.1246
sadness : 0.6979
surprise : 0.0617

*Transformer-based



Happy – Clear Example

Today was amazing. I woke up feeling refreshed for once, and everything just flowed naturally. I got an email saying I'd been selected for a program I'd been dreaming about for months. My friends surprised me with a mini celebration, and we laughed until our cheeks hurt. I felt seen, supported, and genuinely alive — like everything I've been working for is finally coming together.

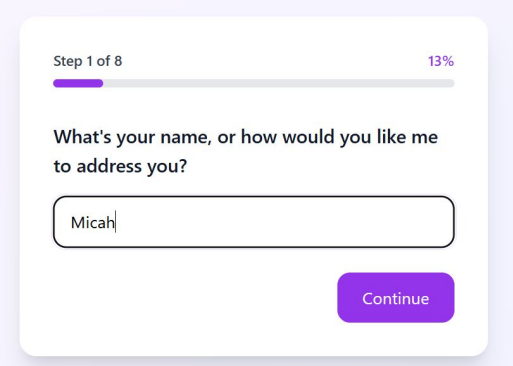
Happy – Nuanced Example

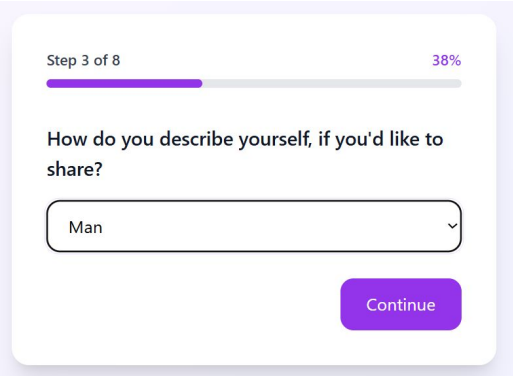
Sad – Clear Example

Sad – Nuanced Example

Analyzer: About User

Initial Form





About User ANALYSIS

about_user.txt

```
{
  "certain": {
    "The user values joy and friendship.": true,
    "The user feels confident in their abilities.": true
  },
  "unsure": {
    "The user may have recently formed new social connections.":
      "short_term": 0.7,
      "long_term": 0.5
    },
```

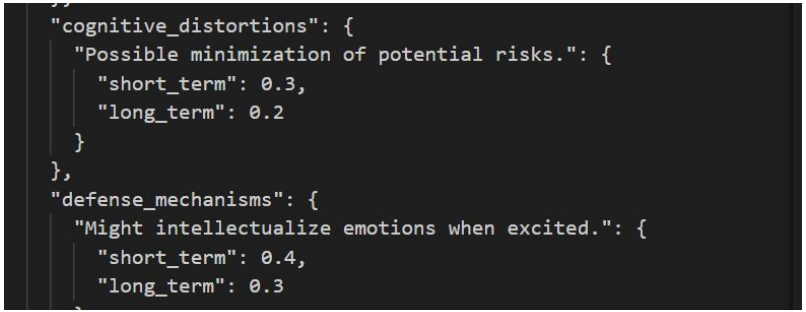
The chatbat **stores certain and uncertain information about the user with respective probabilities**

- This will be used to give a more empathetic interaction between the model and user

Analyzer: Psychoanalysis (How does a therapist think?)

To form an intimate relationship with the user, the model **thinks deeply** about the user and adds its thoughts to the context window

- Long and short term probabilities included



🧠 1. Cognitive Distortions 🪞 2. Defense Mechanisms 📈 3. Maladaptive Patterns 🧩 4. Inferred Beliefs / Self-Schema 📊 5. Emotional Regulation Patterns

- good or bad thinking patterns - axioms that the user believes about the world or themself. core beliefs -Cognitive Distortions -2. Defense Mechanisms - 3. Maladaptive Patterns - Inferred

Responder

- Responds based on all previous context
- Takes into consideration all aspects of user with nuanced context window
- Passes result to the frontend after running in the backend