Emotion detection in Conversational AI

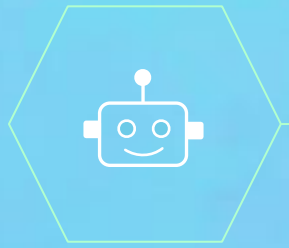
Steering Lab by Horváth & Partners GmbH

Mehdi Moussaoui, Jiaqi Wang, Carolin Bauerhenne,
Alaa Gammoudi, Jessica Reichelt

Mentors: Olena Schüssler, Inna Vasylchuk,
& Michael Rauchensteiner

Content

Overview



Task

Identify User Emotions while chatting
→ improve human-like behaviour
in digital assistants

Motivation

More personal conversations
Better Chatbots
Higher customer satisfaction

Plan

Train a Chatbot on suitable data to
identify emotions from chatting

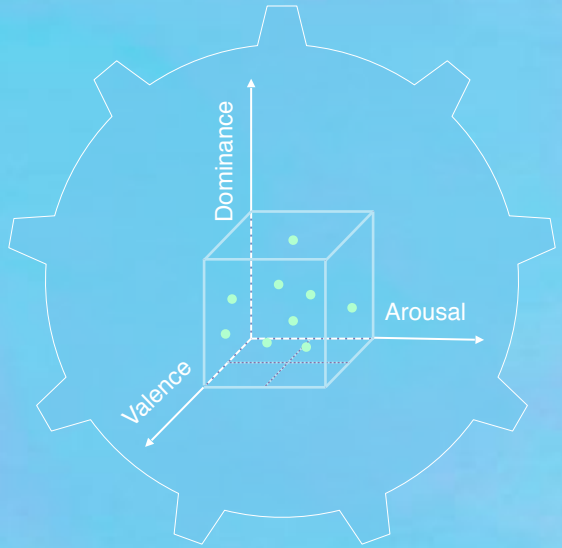
Emotion Models

Categorical models

Ekman's Six Basic Emotions

Dimensional models

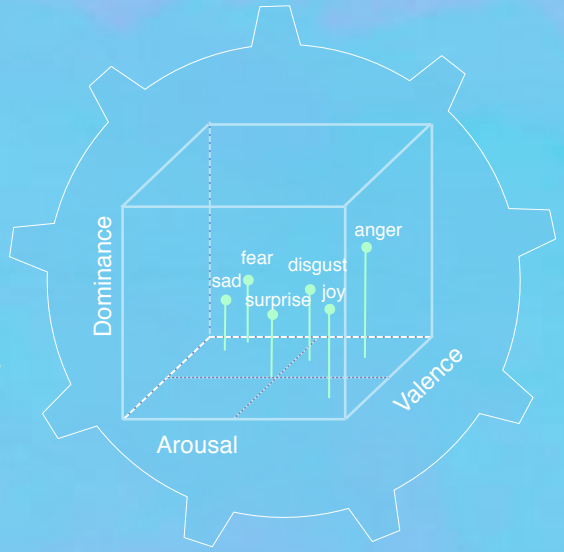
VAD



Emotion Models

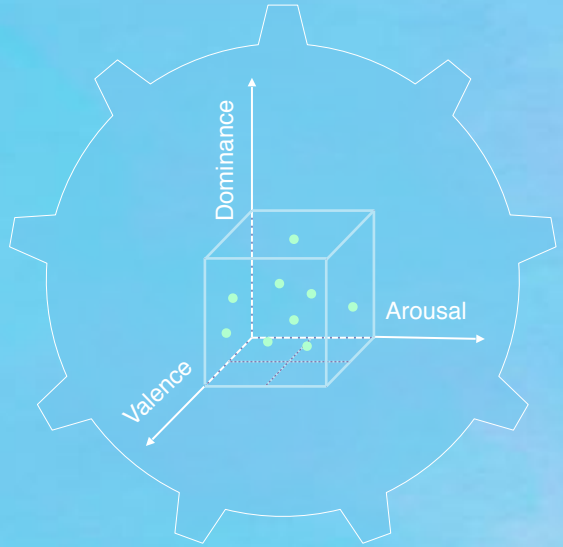
Categorical models

Ekman's Six Basic Emotions



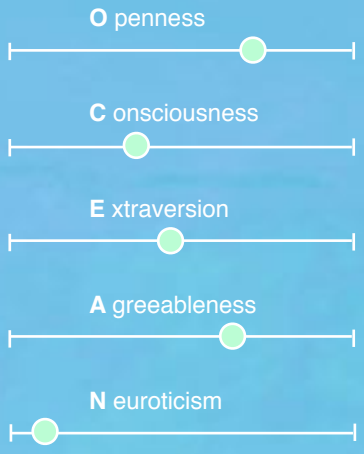
Dimensional models

VAD



Personalized reactions

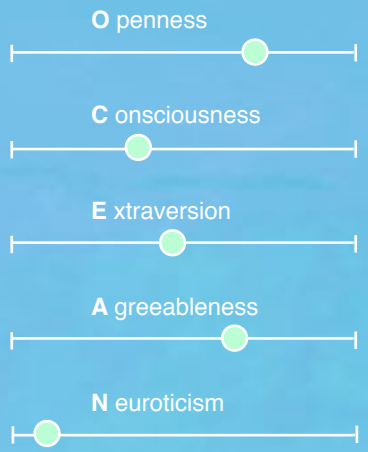
Personality

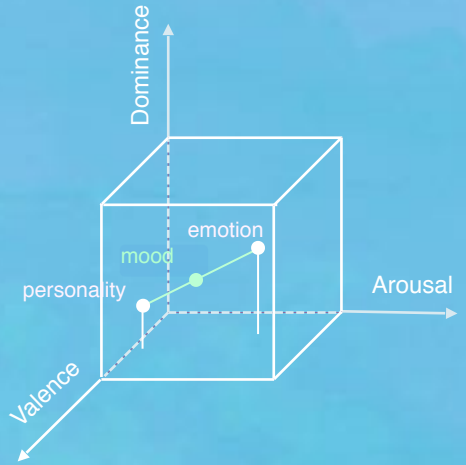


Personalized reactions

Personality

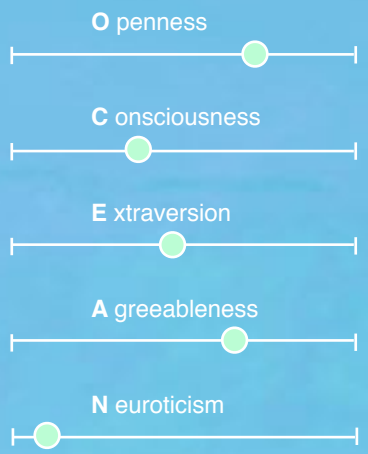
Mood

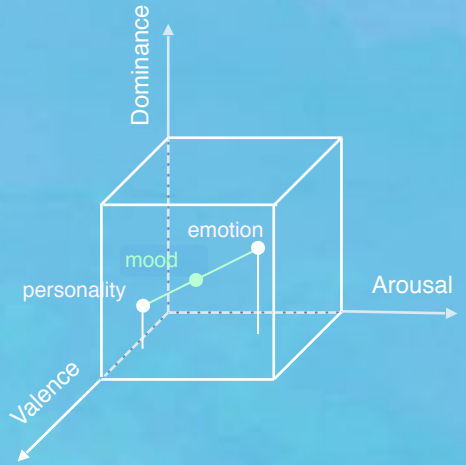


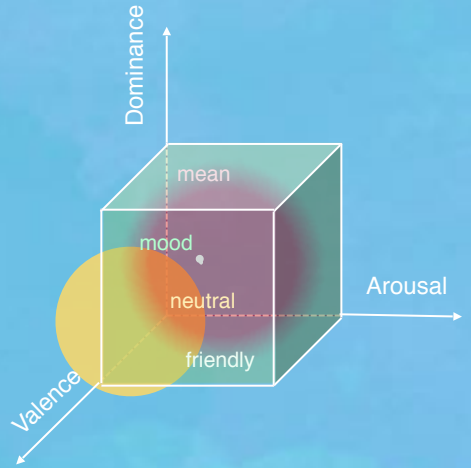


Personalized reactions

Personality → Mood → Reaction







Dataset Selection

Criteria	Description
Emotion model	categorical vs dimensional
Granularity & topic:	language style and data type
	+ : posts on social media, dialogues
	- : news, literatures, headlines
Size	sufficient data and balanced labels
Label	single or multi label

Dataset Selection

SemEval:

Multi-label dataset

11 emotion categories

Anti.	Love	Opti.	Pessi.	Trust
13.9%	12.3%	31.3%	11.6%	5.0%

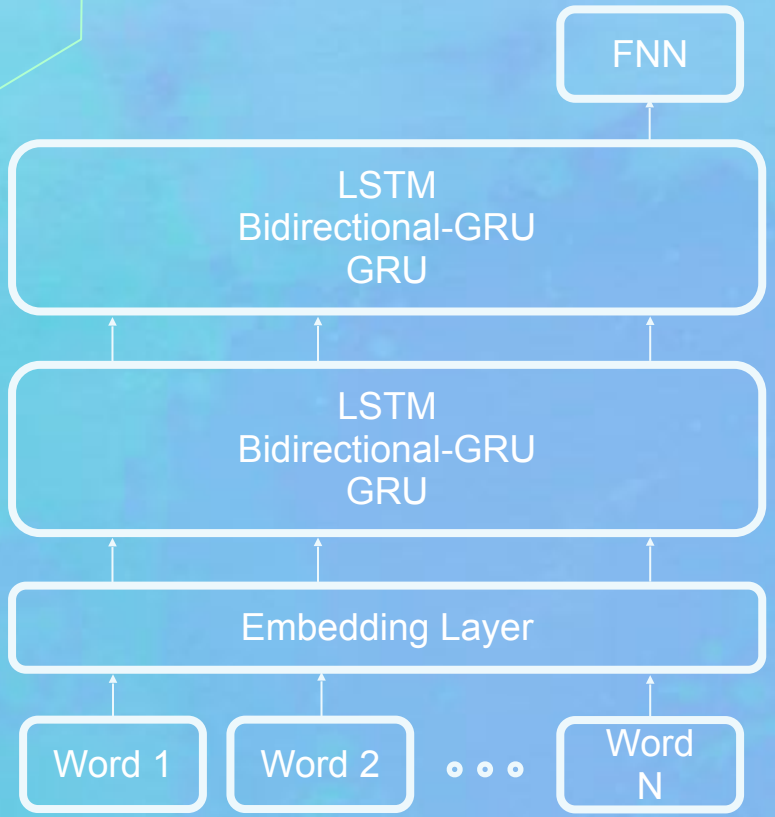
Unified Dataset:

12 merged unified datasets

Single label

Dataset	Joy	Anger	Sadness	Disgust	Fear	Surprise
SemEval	39.3%	36.1%	29.4%	36.6%	16.8%	5.2%
Unified Dataset	36.7 %	10.9 %	19.1 %	6.4 %	17.0 %	9.9 %

Deep Learning Models



SemEval Model:

Outputs 7 classes

and corresponding probability

Unified Dataset Model:

Outputs 1 class

with maximum probability

BERT‘S Architecture

SVM

Single-label binary classification on the SemEval dataset

	Joy		Surprise		Sadness		Anger		Fear		Disgust		Neutral	
Parameters	C = 100 γ = scale		C = 100 γ = auto		C = 100 γ = scale		C = 1 γ = auto		C = 100 γ = scale		C = 1 γ = scale		C = 100 γ = scale	
Accuracy	0.51		0.93		0.66		0.60		0.79		0.61		0.89	
F-score	**0**	**1**	**0**	**1**	**0**	**1**	**0**	**1**	**0**	**1**	**0**	**1**	**0**	**1**
	0.63	0.27	0.97	0.03	0.78	0.17	0.73	0.17	0.88	0.20	0.75	0.20	0.94	0.11

Deep learning models comparison

On SemEval:

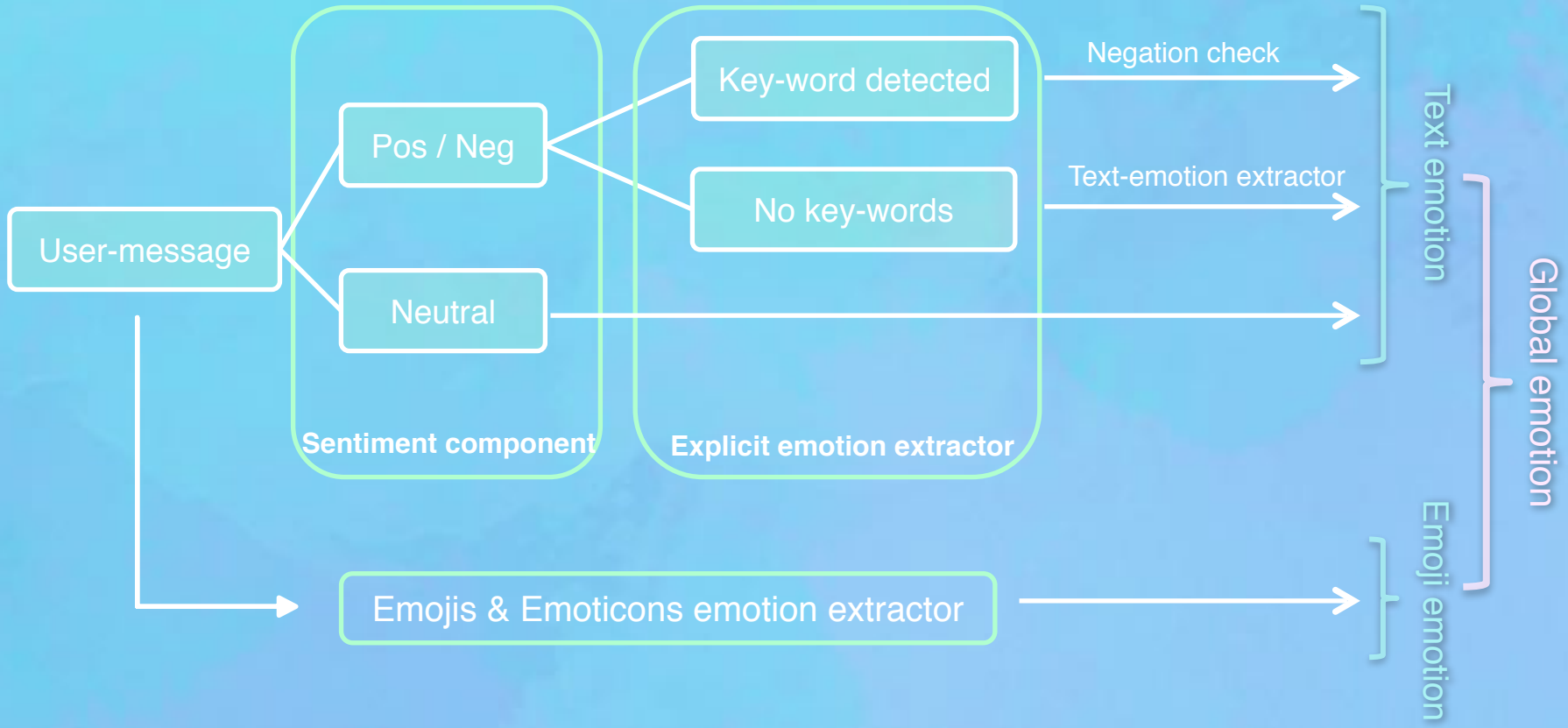
Model (threshold) Dataset	Metrics	Anger	Disgust	Fear	Joy	Sadness	Surprise	Macro Average
RNN-LSTM (0.8) SemEval	*Precision*	0.53	0.50	0.52	0.71	0.70	0.40	0.560
	Recall	0.47	0.55	0.45	0.60	0.57	0.27	0.485
	F-score	0.498	0.524	0.482	0.650	0.628	0.323	0.520
RNN-BiGRU (0.8) SemEval	*Precision*	0.50	0.46	0.48	0.73	0.55	0.42	0.523
	Recall	0.34	0.34	0.35	0.66	0.56	0.32	0.428
	F-score	0.405	0.391	0.405	0.693	0.555	0.363	0.469
RNN-GRU (0.8) SemEval	*Precision*	0.34	0.26	0.38	0.64	0.54	0.37	0.422
	Recall	0.29	0.23	0.37	0.41	0.65	0.34	0.385
	F-score	0.313	0.244	0.375	0.523	0.590	0.354	0.423
BERT (0.8) SemEval	*Precision*	**0.889**	**0.858**	**0.897**	**0.943**	**0.925**	**0.500**	**0.83**
	Recall	0.635	0.571	0.579	0.703	0.419	0.393	0.54
	F-score	0.741	0.686	0.704	0.805	0.577	0.440	0.65
BERT (0.5) SemEval	*Precision*	0.815	0.782	0.754	0.883	0.763	0.508	0.74
	Recall	**0.781**	**0.799**	**0.736**	**0.795**	**0.608**	**0.411**	**0.69**
	F-score	**0.797**	**0.791**	**0.745**	**0.837**	**0.676**	**0.454**	**0.71**

Deep learning models comparison

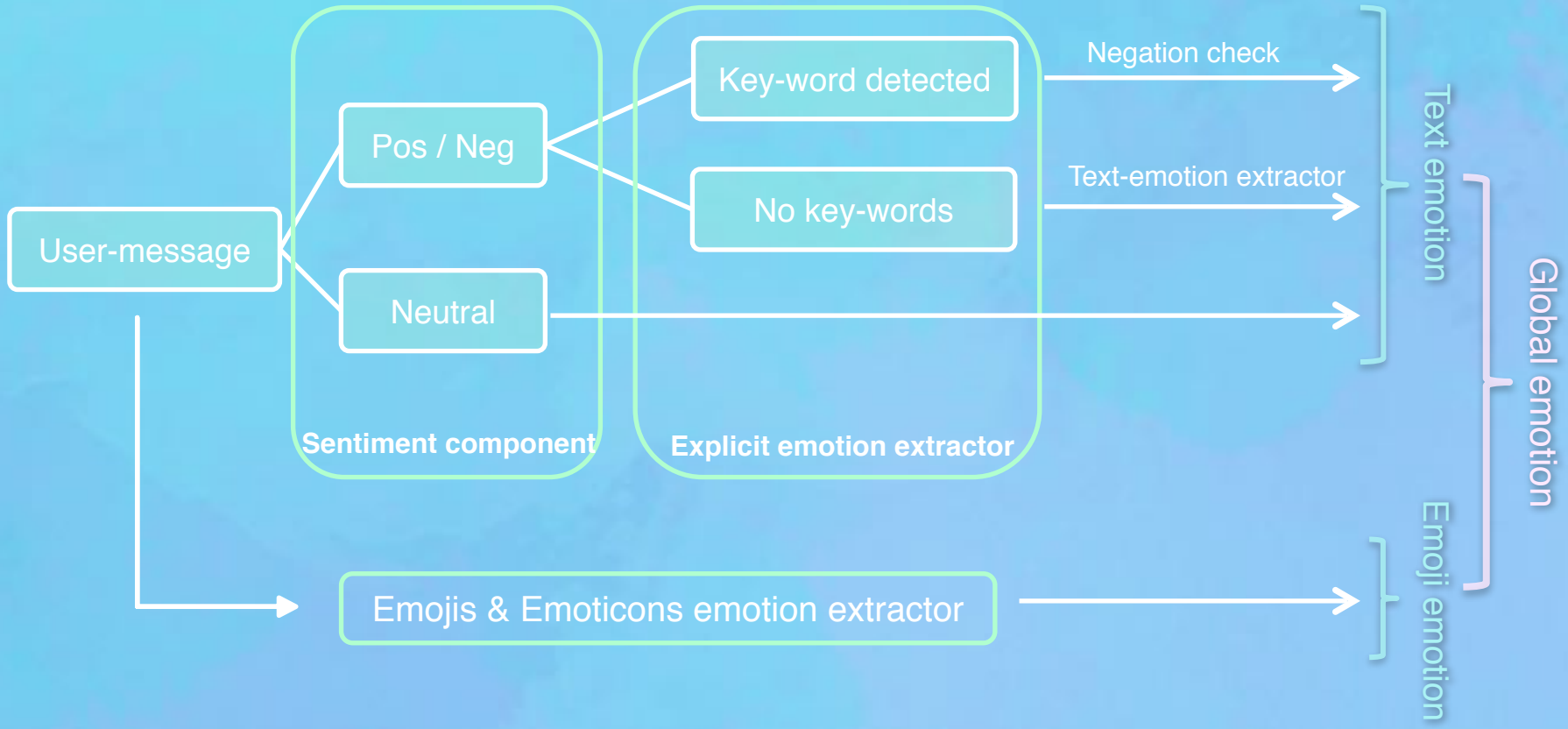
On Unified dataset:

Model (thres.) Dataset	Metrics	Anger	Disgust	Fear	Joy	Sadness	Surprise	Macro Average
RNN-LSTM Unified	*Precision* *Recall* *F-score*	**0.701** **0.476** **0.567**	0.700 **0.472** **0.564**	**0.749** 0.449 0.546	**0.910** **0.532** **0.671**	**0.854** **0.601** **0.706**	**0.533** **0.471** **0.500**	**0.741** **0.500** **0.597**
RNN-BiGRU (0.8) Unified	*Precision* *Recall* *F-score*	0.700 0.432 0.534	0.691 0.402 0.508	0.692 **0.452** **0.547**	0.859 0.500 0.632	0.822 0.528 0.643	0.492 0.453 0.472	0.709 0.461 0.558
RNN-GRU (0.8) Unified	*Precision* *Recall* *F-score*	0.690 0.380 0.490	**0.703** 0.374 0.488	0.730 0.423 0.536	0.906 0.441 0.593	0.845 0.477 0.610	0.542 0.403 0.462	0.736 0.416 0.532

Hierarchical emotion detection



Hierarchical emotion detection



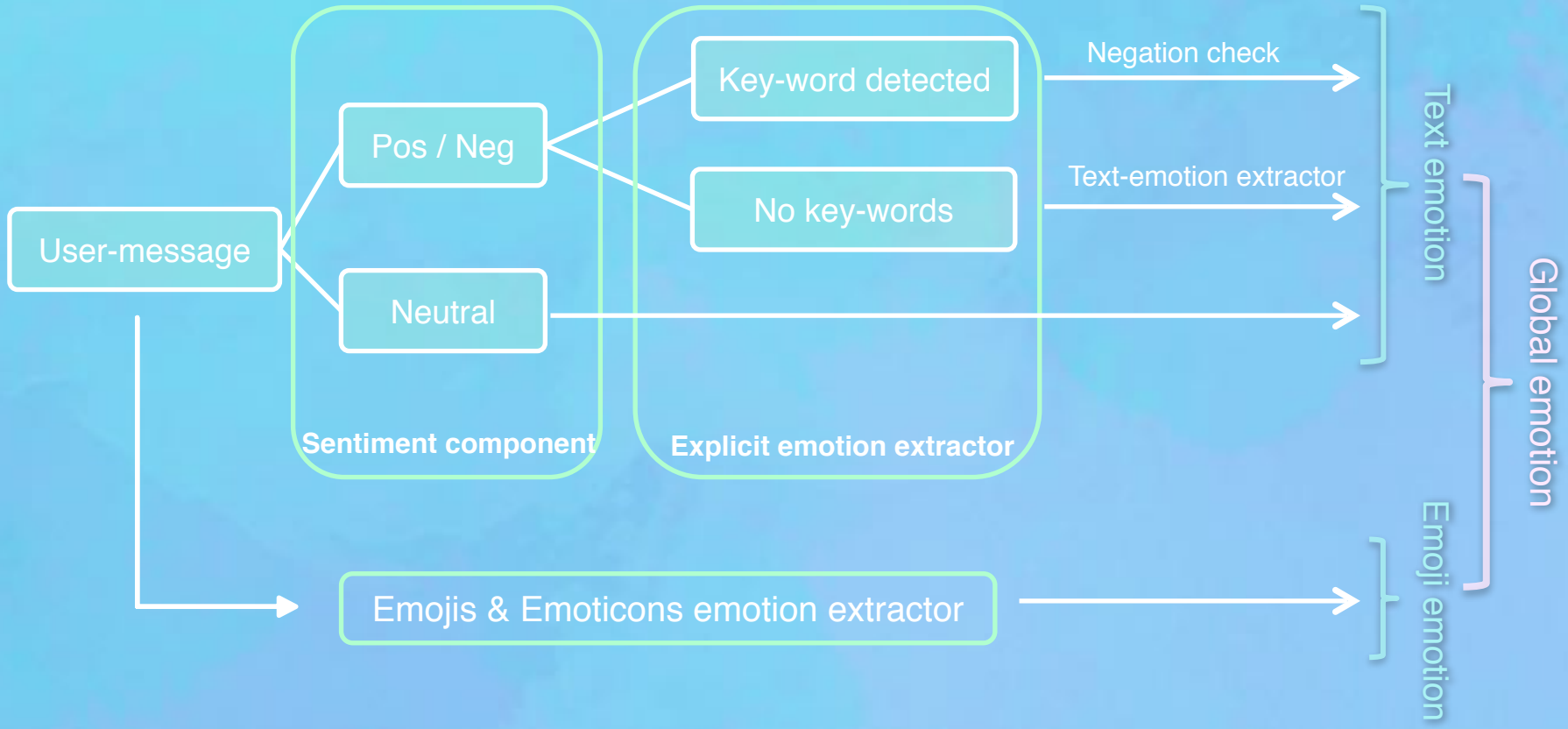
Sentiment Analysis

- Detects polarity within text: *positive, negative, or neutral*
- Neutral sentiment → neutral emotion
- Pos. or neg. sentiment → prediction with emotion extraction models

✓ Reduced complexity and computational effort

✓ Increased reliability of the neutral emotion

Hierarchical emotion detection

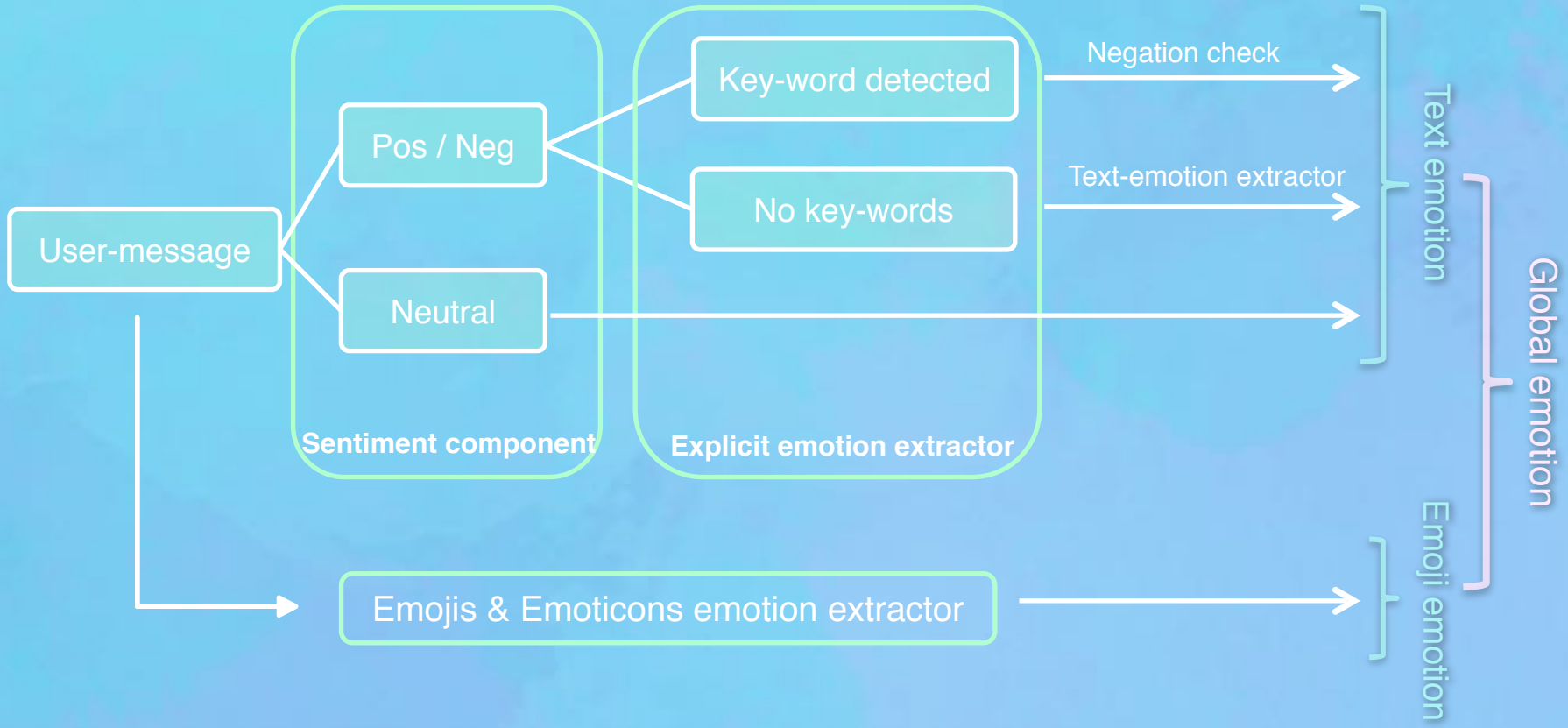


Key-word approach

- Explicit emotions directly recognized if key-words are present
    1. Detect key-words
    2. Negation check

- Examples:
    - *“I am happy with your service”* → key-word happy → emotion : *joy*
    - *“The drink is not bad!”* → key-word bad + negation not → emotion : *joy*

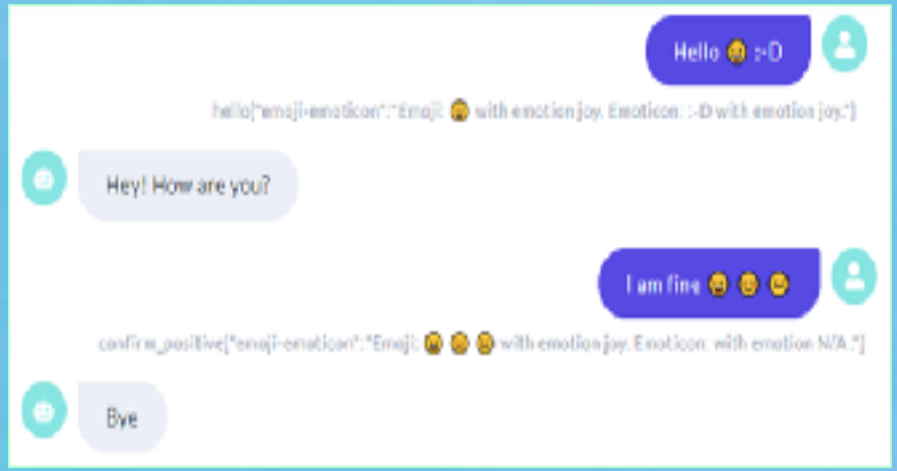
Hierarchical emotion detection



Emoji & Emoticon Detection

- ✓ Pictorial depiction of the facial expression
- ✓ Language-independent indicators of emotions

- ⬡ Building an Emoticon & Emoji database
- ⬡ Emotion extraction based on frequency

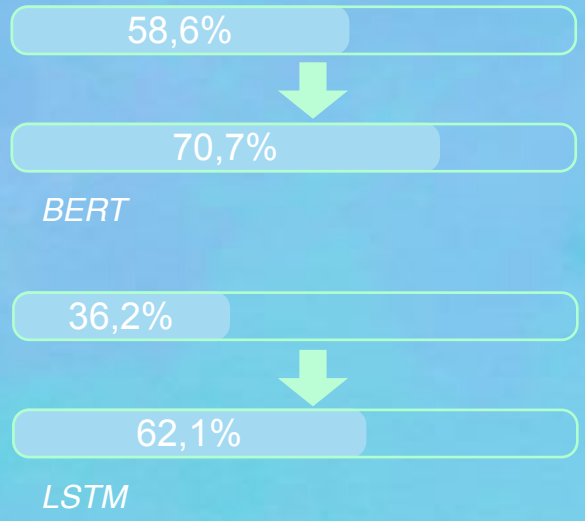


Preprocessing Results

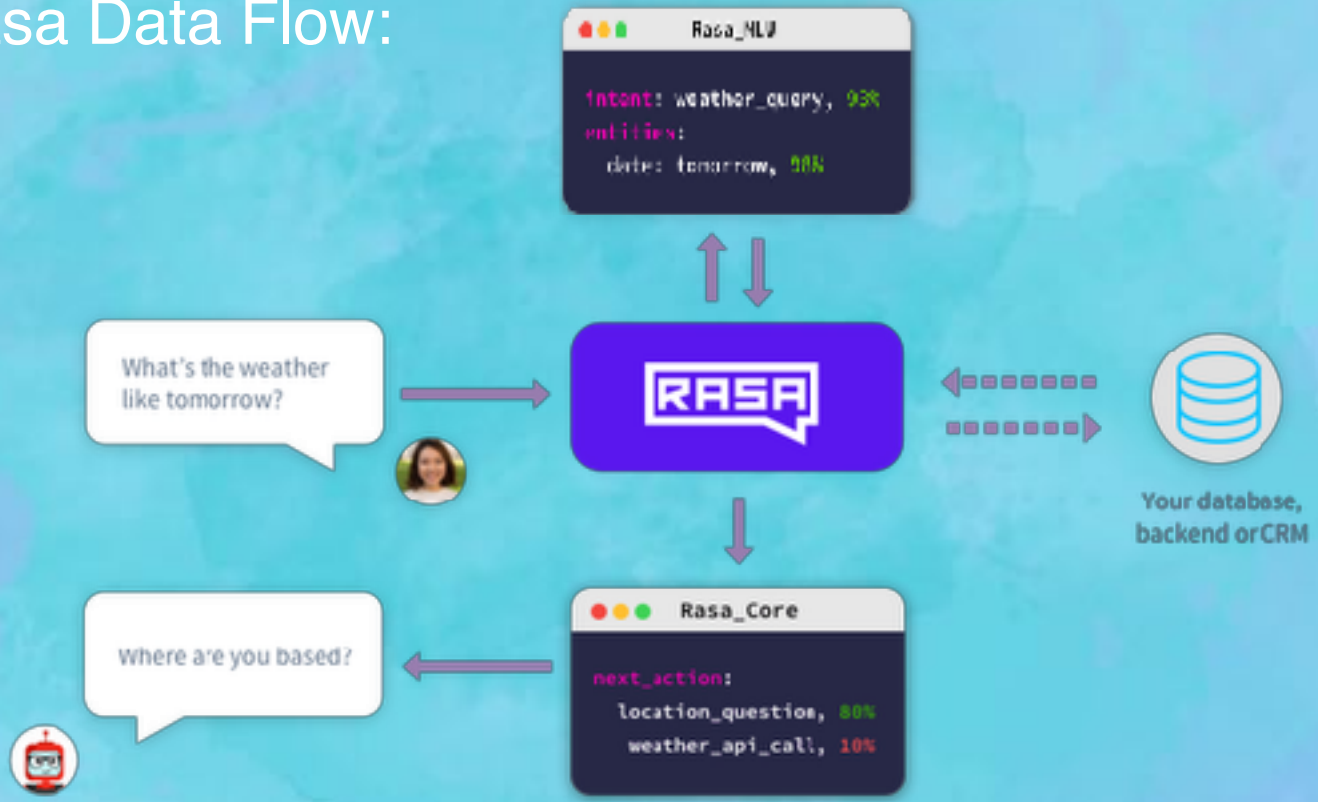
Advantages:

- Faster estimations of the emotions
- Better accuracy on Feedback sentences
- Sarcasm Detection using Emojis / Emoticons

Accuracy Results:

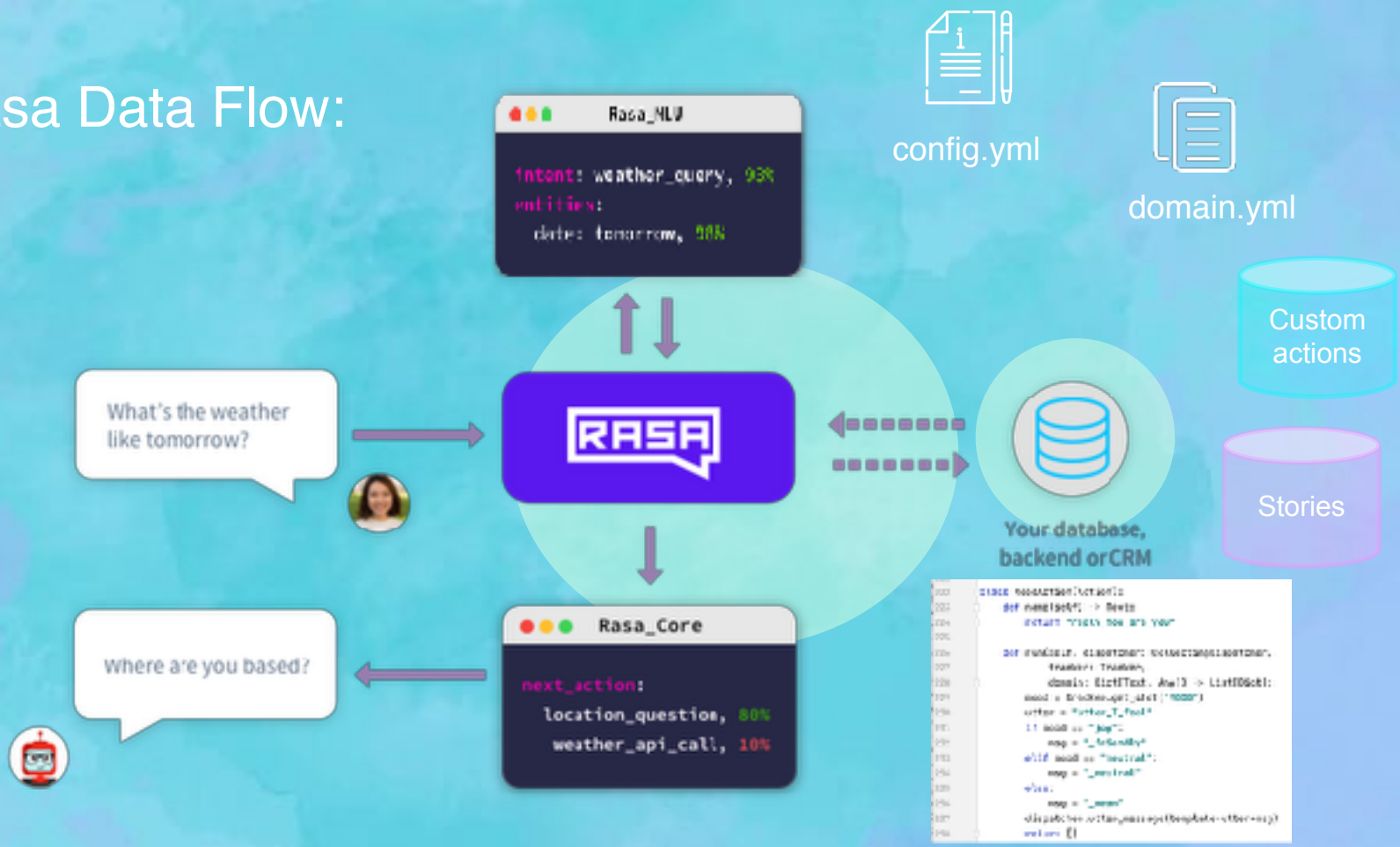


Rasa Data Flow:



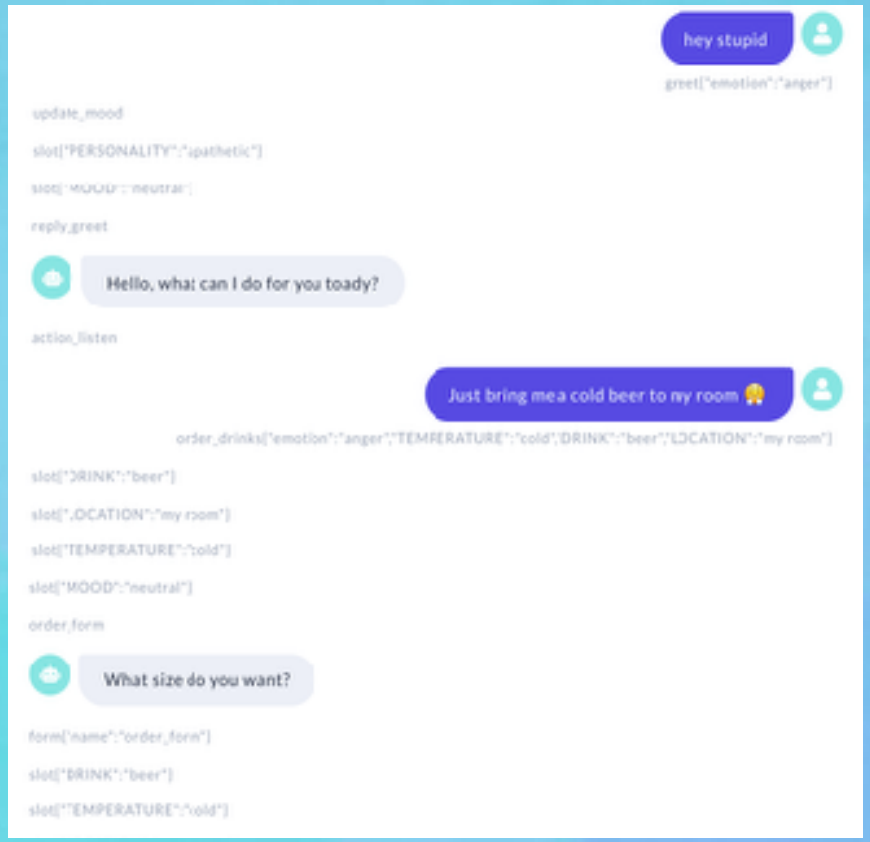
https://chatbotslife.com/building-a-weather-chatbot-with-rasa-nlu-and-rasa-core-8a295f29c405a

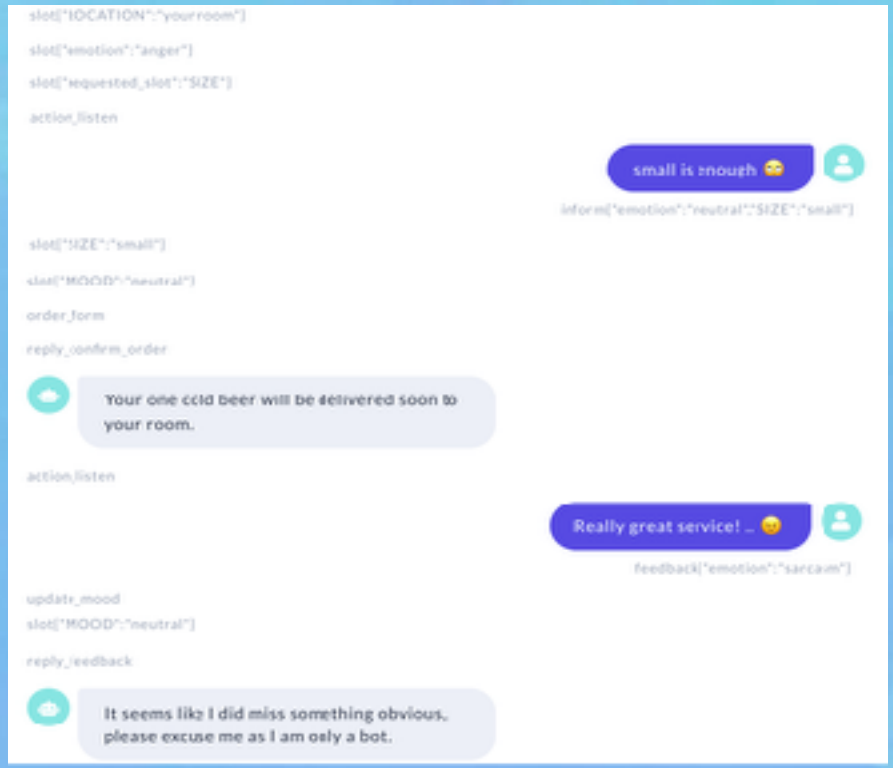
Rasa Data Flow:



https://chatbotslife.com/building-a-weather-chatbot-with-rasa-nlu-and-rasa-core-8a295f29c405a

Example


hey stupid
greet{"emotion":"anger"}
update_mood
slot{"PERSONALITY":"apathetic"}
reply_greet
Hello, what can I do for you toady?
action_listen
Just bring me a cold beer to my room
order_drinks{"emotion":"anger","TEMPERATURE":"cold","DRINK":"beer","LOCATION":"my room"}
slot{"DRINK":"beer"}
slot{"LOCATION":"my room"}
slot{"TEMPERATURE":"cold"}
slot{"MOOD":"neutral"}
order_form
What size do you want?
form{"name":"order_form"}
slot{"DRINK":"beer"}
slot{"TEMPERATURE":"cold"}


slot{"LOCATION":"your room"}
slot{"emotion":"anger"}
slot{"requested_slot":"SIZE"}
action_listen
small is enough
inform{"emotion":"neutral","SIZE":"small"}
slot{"SIZE":"small"}
slot{"MOOD":"neutral"}
order_form
reply_confirm_order
Your one cold beer will be delivered soon to your room.
action_listen
Really great service! ..
feedback{"emotion":"sarcasm"}
update_mood
slot{"MOOD":"neutral"}
reply_feedback
It seems like I did miss something obvious, please excuse me as I am only a bot.

Conclusion

Project Goal

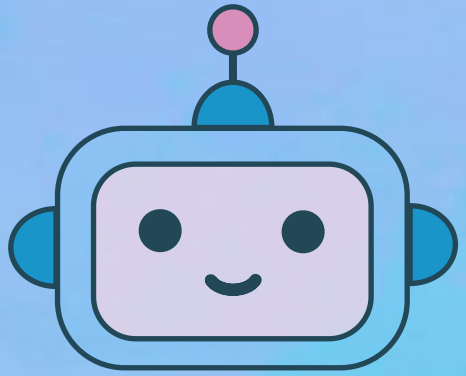
Affective personal assistant
Ability to have own sentiments, emotional reactions and a personality
Acceptable accuracy

Adaptability

Our emotion extraction model can easily be adapted to different psychological approaches and platforms.

Outlook

Add other personalities
Expand action-/NLU-/stories- file.
Integrate speech and image recognition
Extend the chat bot to other service domains

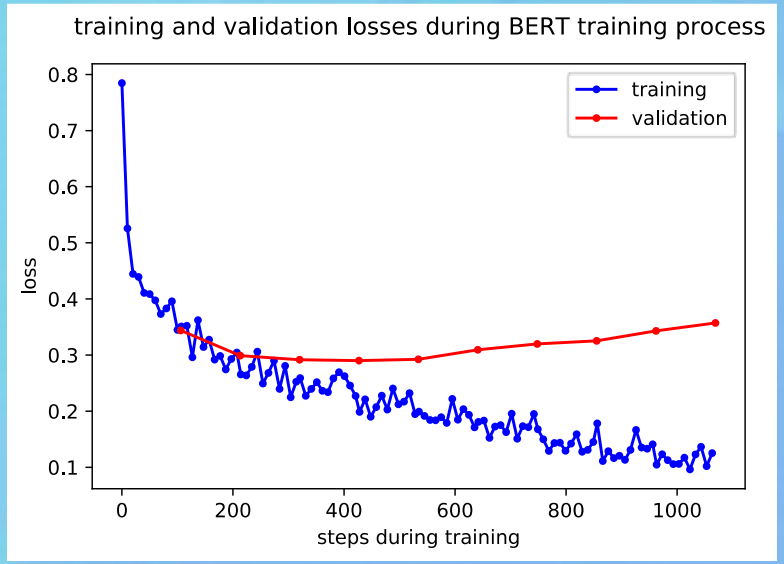


Thank you!

CREDITS: This presentation template was created by Slidesgo, including icons by Flaticon, and infographics & images by Freepik.

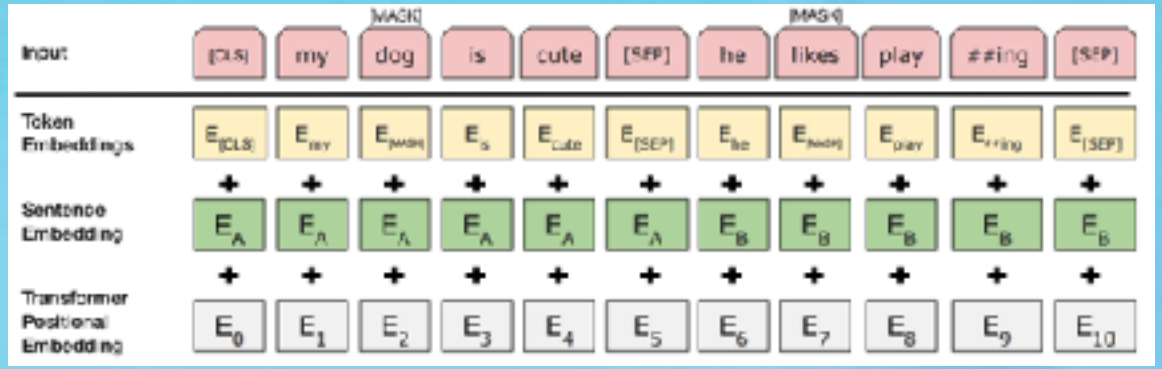
Training BERT*

Hyperparameters:	Learning rate	Weight-decay	Batch size	Epoch
Values:	2e-05	3e-06	64	10

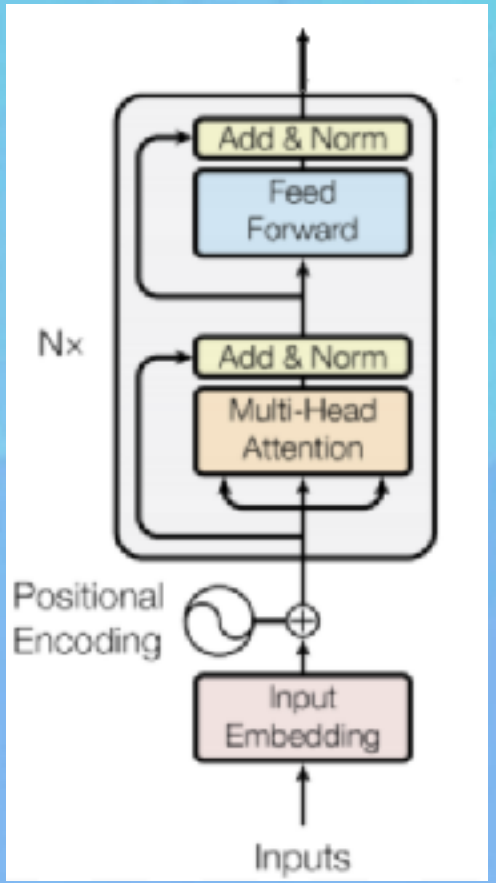


* On SemEval

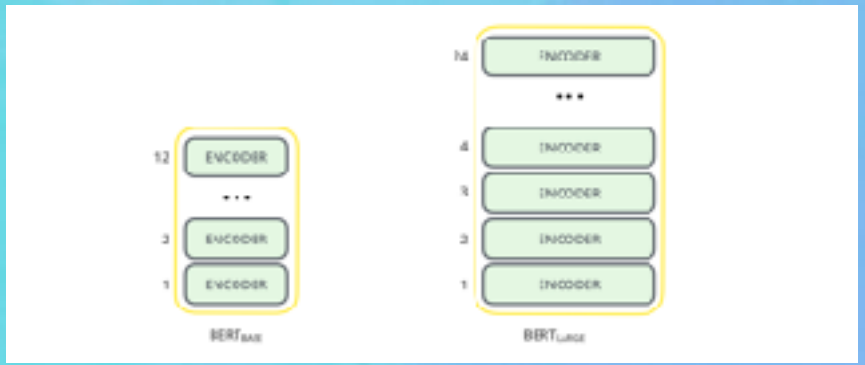
Inputs representation



Encoders' architecture



BERT's architecture

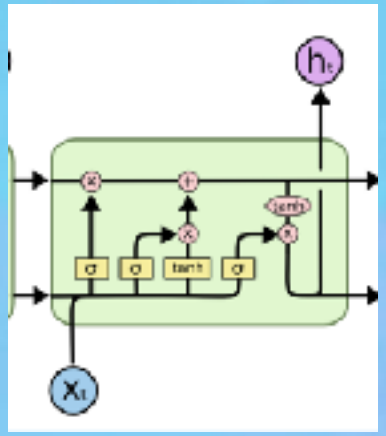


BackUp – Dataset

Dataset	Granularity	Annotation	Size	Topic	Label
Affective Text	Headline	E + V	1,250	news	multi
Blogs	Sentence	E + ne + me	5,025	blogs	single
CrowdFlower	Tweet	E + CF	40,000	general	single
DailyDialog	Dialogue	E	13,118	multiple	single
EmoBank	Sentence	V + A + D	10,548	multiple	single
ISEAR	Description	E + SG	7,665	events	single
SSEC	Tweet	P	4,868	general	multi
Gb-valence-arousal	Facebook	V + A	2,895	questionnaire	single
SemEvel-2018(Task 1)	Tweet	P + M	7,902	general	multi

[E] Ekman: anger, disgust, fear, joy, sadness, surprise; [V] valence; [A] arousal; [D] dominance; [ne] no emotion; [me] mixed motion; [SG] shame, guilt; [CF] enthusiasm, fun, hate, neutral, love, boredom, relief, empty; [P] anger, disgust, fear, joy, sadness; [M] love, optimism, pessimism

LSTM & GRU



Goal:

Avoid gradient vanishing

LSTM:

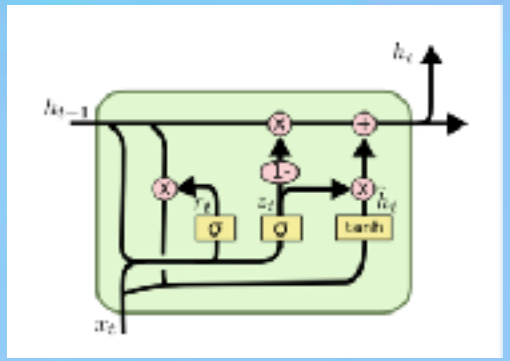
forget gate: $f_t = \sigma\left(W_f \cdot \left[h_{t-1},\ x_t\right] + b_f\right)$

input gate: $i_t = \sigma\left(W_i \cdot \left[h_{t-1},\ x_t\right] + b_i\right)$

output gate: $o_t = \sigma\left(W_o \cdot \left[h_{t-1},\ x_t\right] + b_o\right)$

cell state: $C_t = f_t\ * C_{t-1} + i_i\ *\ \tilde{C}_t$

hidden state: $h_t = o_t\ * \tanh(C_t)$



GRU:

update gate: $z_t = \sigma\left(W_z \cdot \left[h_{t-1},\ x_t\right]\right)$

reset gate: $r_t = \sigma\left(W_r \cdot \left[h_{t-1},\ x_t\right]\right)$

source: https://colah.github.io/posts/2015-08-Understanding-LSTMs/

Hierarchical Emotion Detection - Components

- Emoji and Emoticon
- Sentiment Analysis
- Key-word approach
- Fusion of text and Emoji/Emoticons detections

Rasa Custom Actions