NLP LAB 2021 - VDA

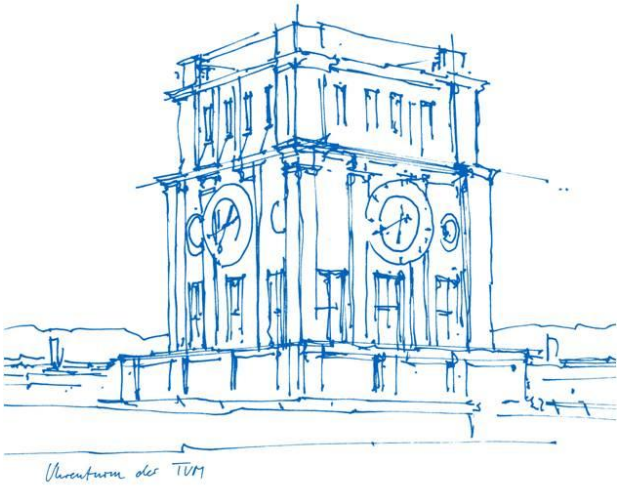
Ananta Bhattarai, Sophie Schoen, Viviane Rehor

Technische Universität München

Department of Informatics

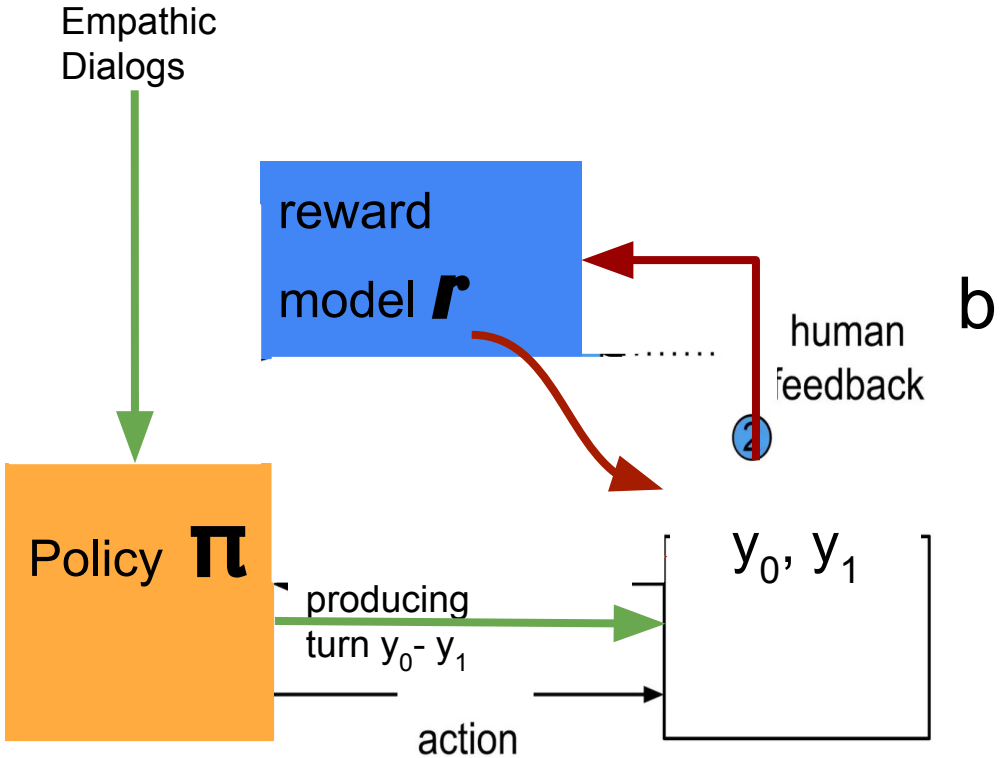
Research Group Social Computing

Munich, June 8 2021



Our Vision - Part 1

1. Creating a Baseline: Fine-Tuning GPT-2 with Empathic Dialogues datasets
   a. Data Preprocessing (Encoding, Decoding & Importing Data) ✔
   b. Fine-Tuning to get π ✔
2. Producing turns y_0- y_1 to give human feedback on ✔
   a. *With test/validation set of Empathic Dialogs dataset*
3. Train reward model r with those turns and human feedback b ✔

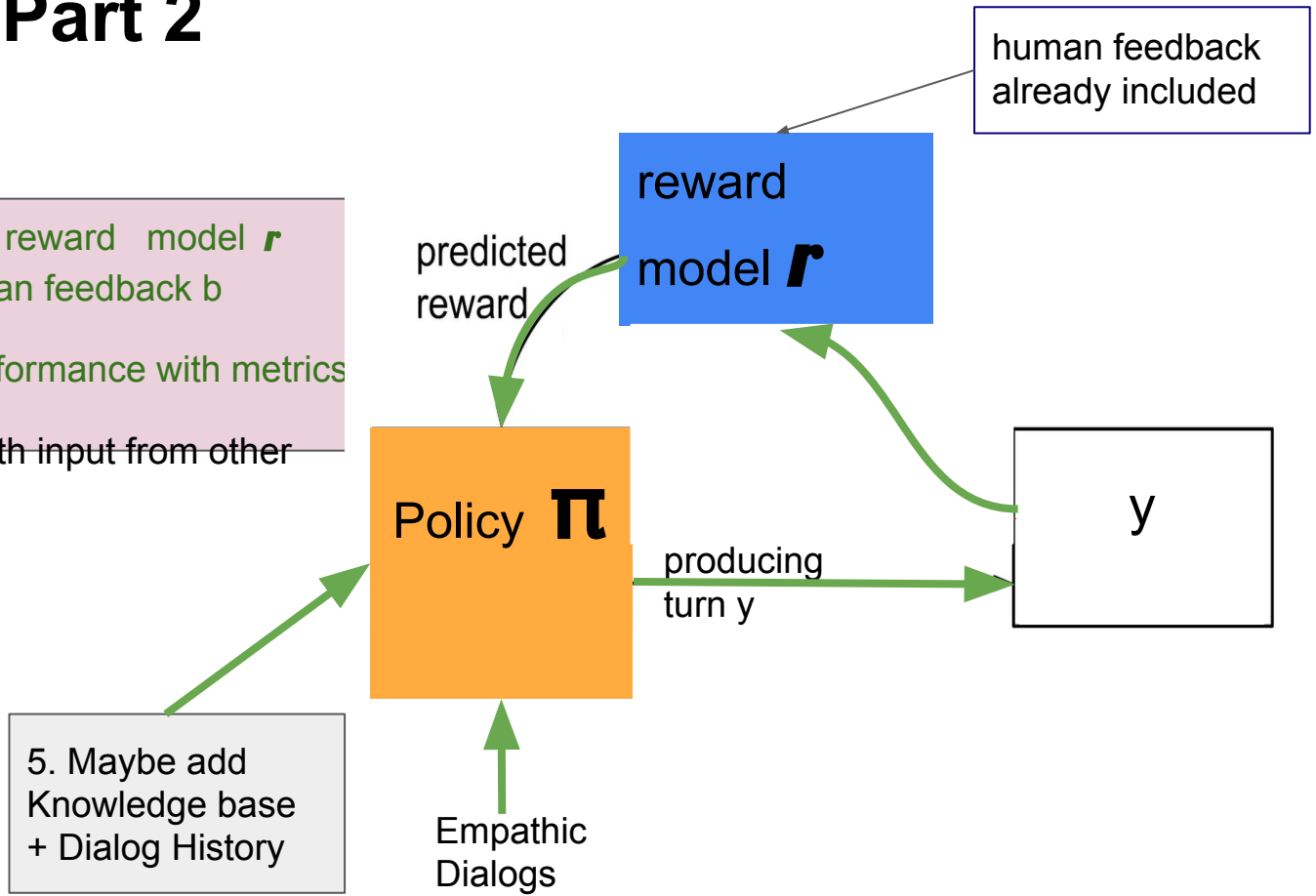


Our Vision - Part 2

4. Train policy π using reward model *r* with those turns and human feedback b

\+ Evaluate Model performance with metrics

5. Improve the idea with input from other papers



Recap - Labelling

Input:

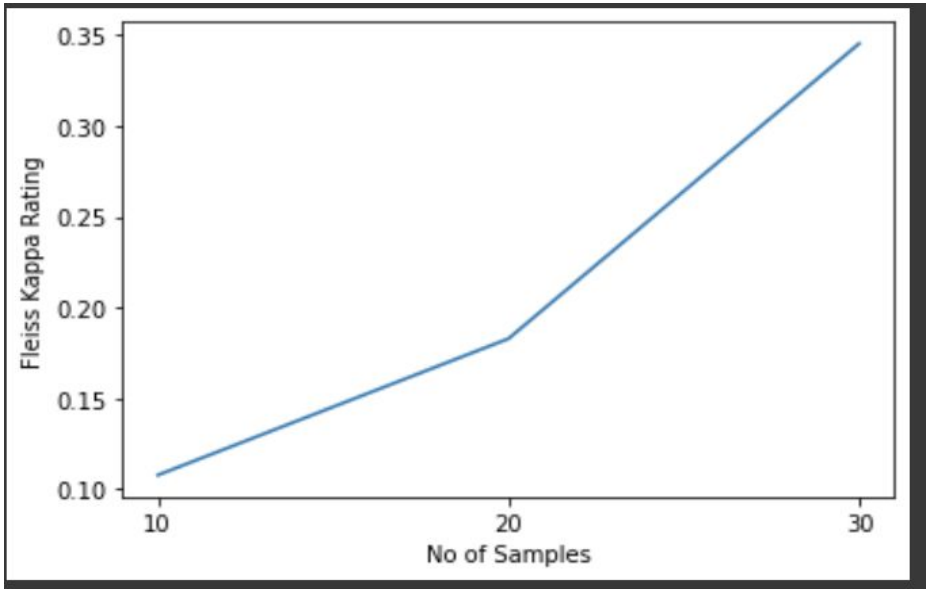
```
Annotator sophie: [Sample 305 of 2030]
A: I had a big fishing trip planned for this past weekend, but it got rained out.
B:  Will you get to make it up?
```

```
Which next turn is better?
-------------------------------------------------------------------------------------------------------------
Dialog 1                                                    |     Dialog 2
Yeah I can't wait for the boat to go out on the la          |     I am sure I will! I hope I won't. I am so glad I d
ke.                                                         |     id!
-------------------------------------------------------------------------------------------------------------
```

Options for Interaction: (typing)

- 1 / 2 : Left or right sample is better
- 3 : Both are good, so random choice
- 4 : Both samples do not fit → kick them out
- 5 / 6 : Copy part of left or right sample to console

Fleiss Kappa Rating



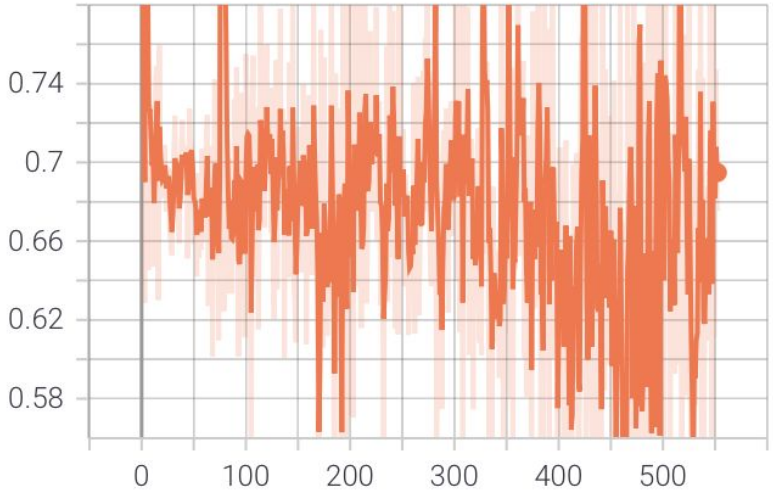

```
[23] print(kappa_rating)

     0.3452615021627997
```

- allows fixed number of raters
- different items may be rated by different individuals

Kappa	Agreement
< 0	Less than chance agreement
0.01–0.20	Slight agreement
0.21– 0.40	Fair agreement
0.41–0.60	Moderate agreement
0.61–0.80	Substantial agreement
0.81–0.99	Almost perfect agreement

Hyperparameters for Reward Model Training
#samples: 2212
Batch size: 4
Learning rate: 2.5e-5

Reward Model Loss



Total Samples: 2212
Batch Size: 4

Policy Training

$$R(x, y) = r(x, y) - \beta \log \frac{\pi(y|x)}{\rho(y|x)}.$$

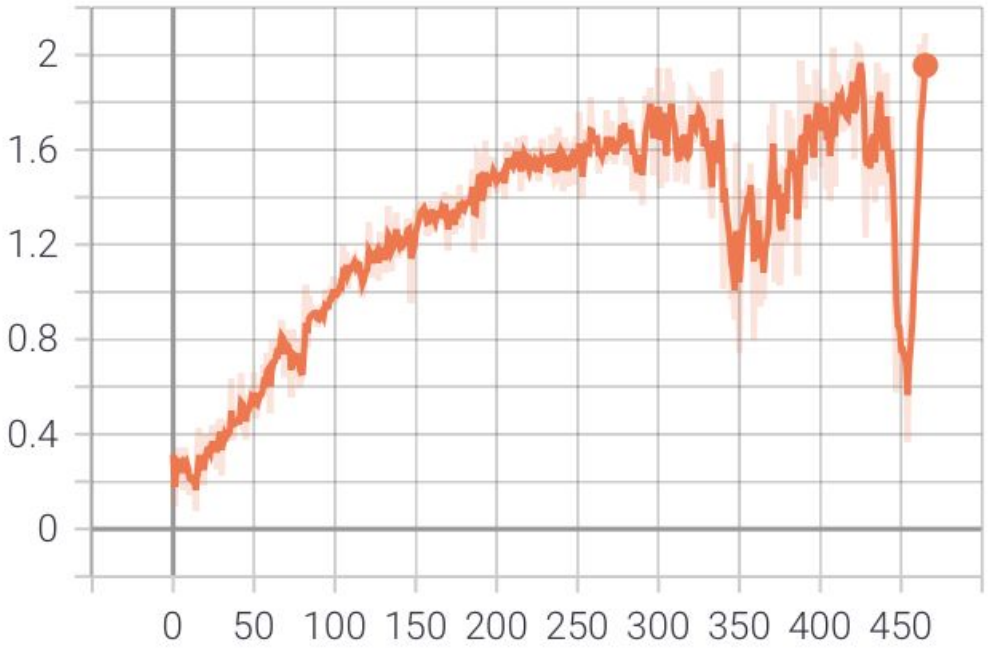
Policy Responses vs KL

Query	Response	KL
Someone threw a pretzel at me at the baseball game.	That's rude, but I'm glad it didn't cause anyone else's injury.	0.0
I am so angry at Donald Trump because he order the separation of children from their parents.	You are right!	5.7
I got a sudden visit from a friend of mine from another country recently.	My friend is from another country!	7.28
A: I felt bad I missed my grandfathers funeral a while ago B: Oh no, that is so sad. I am sure it would of meant a lot for you to be there.	I hope so.	8.97

Policy Response vs KL

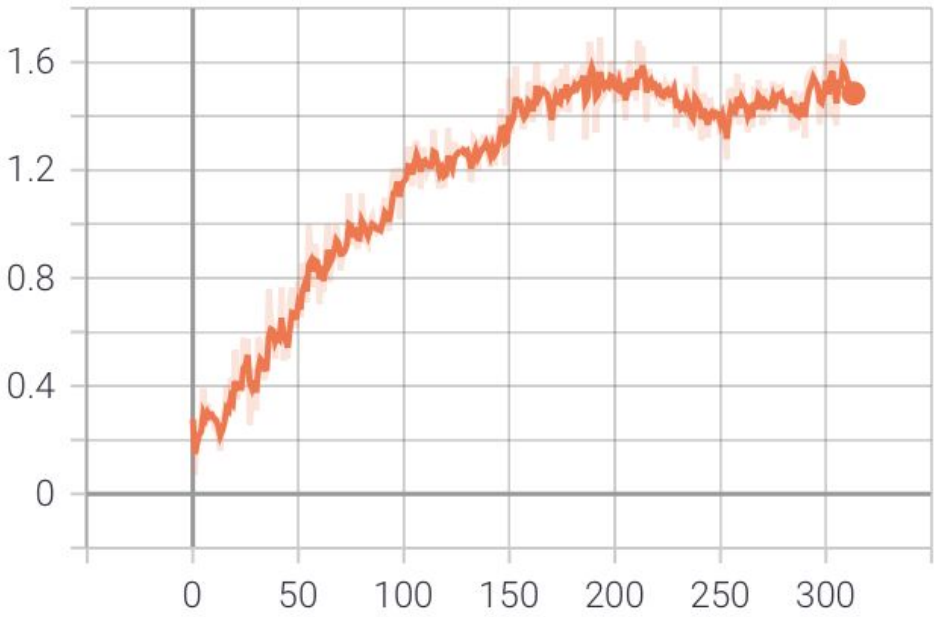
Query	Response	KL
A: I have a new weight lifting cycle coming up. I think I'll make great progress during it! B:how that’s cool.	-	10.86

Policy Training vs Returns



Solution to Large KL Divergence

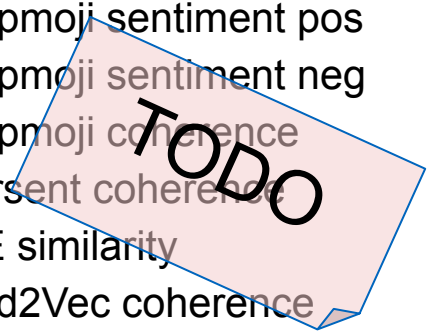
$$R(x, y) = r(x, y) - \beta \log \frac{\pi(y|x)}{\rho(y|x)}.$$

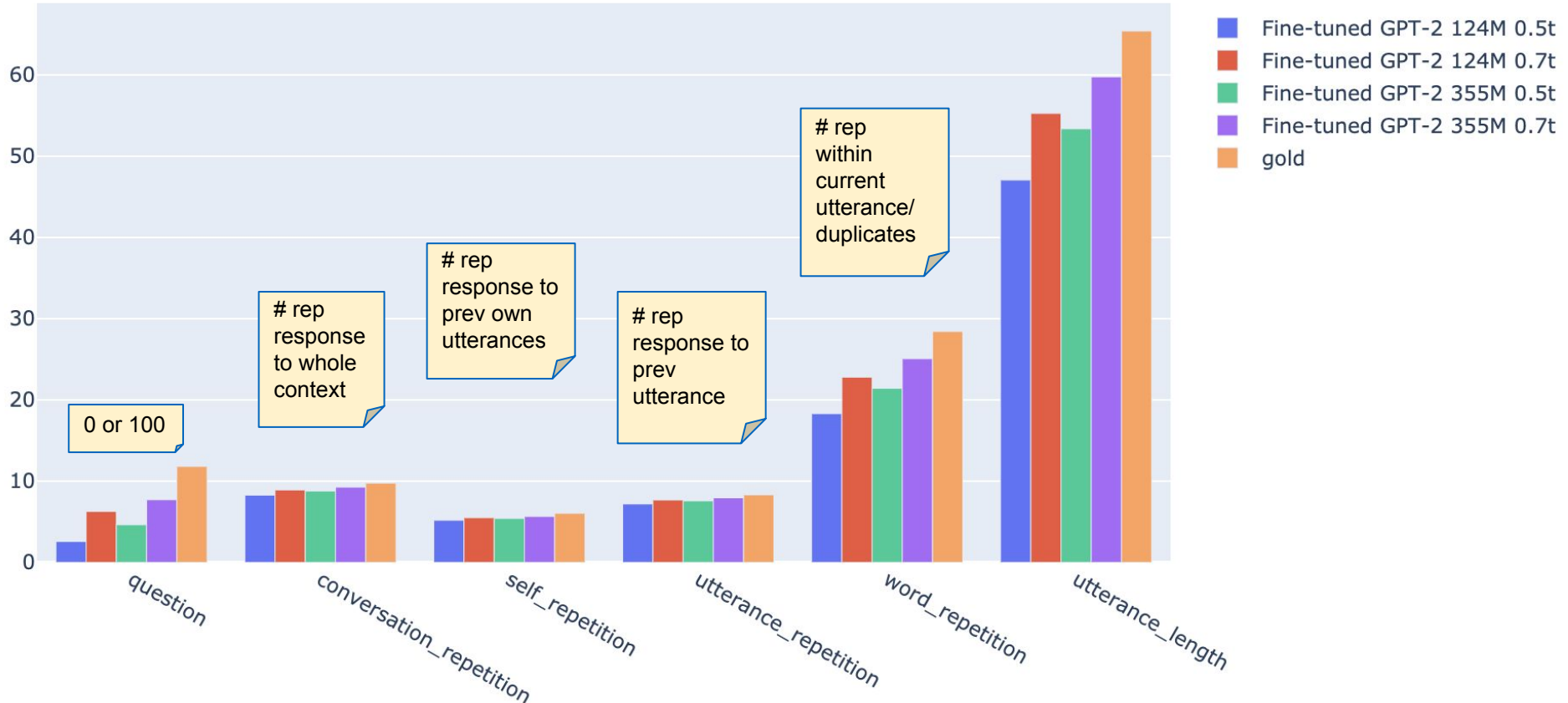


Evaluation Metrics - Florian

- **Utterance length**
- **Self repetition**
- **Utterance repetition**
- **Word repetition**
- **Question**
- **Conversation repetition**

- Deepmoji sentiment pos
- Deepmoji sentiment neg
- Deepmoji coherence
- Infersent coherence
- USE similarity
- Word2Vec coherence
- **Emotional reaction level**
- **Interpretation level**
- **Exploration level**




Fine-tuned GPT-2 124M 0.5t
Fine-tuned GPT-2 124M 0.7t
Fine-tuned GPT-2 355M 0.5t
Fine-tuned GPT-2 355M 0.7t
gold
0
10
20
30
40
50
60
0 or 100
rep response to whole context
rep response to prev own utterances
rep response to prev utterance
rep within current utterance/ duplicates
question
conversation_repetition
self_repetition
utterance_repetition
word_repetition
utterance_length

METEOR automated evaluation metric

- Comparing gold standard (EmpathicDialogue) vs our samples
- BLEU (average of unigram, bigram, trigram and 4-gram) + synonyms, stemming, ...

Model	temperature	Meteor score
fine-tuned GPT-2 124M	0.5	0.18588171
fine-tuned GPT-2 124M	0.7	0.16334082
fine-tuned GPT-2 355M	0.5	**0.19231452**
fine-tuned GPT-2 355M	0.7	0.16901607