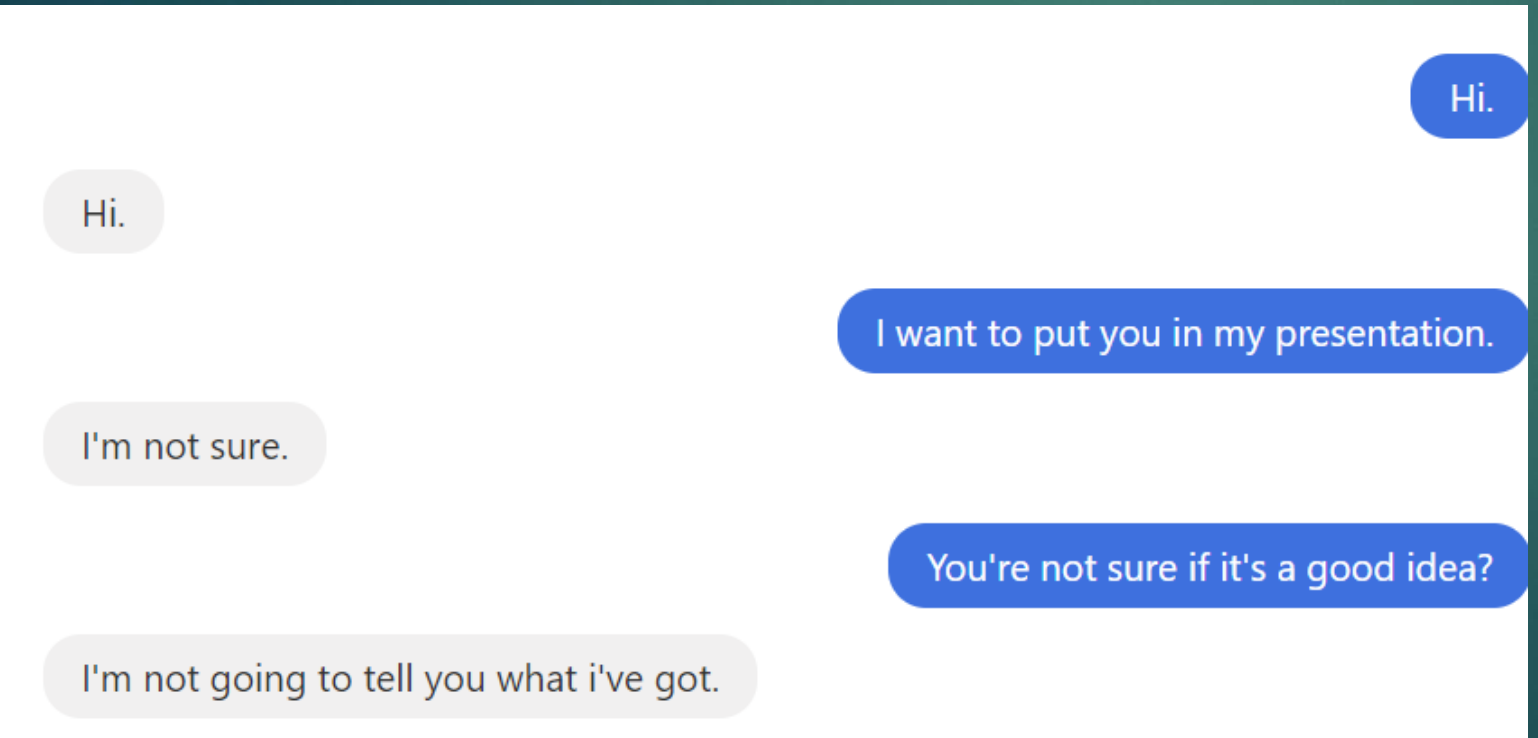
Neural Chatbots

CSÁKY RICHÁRD

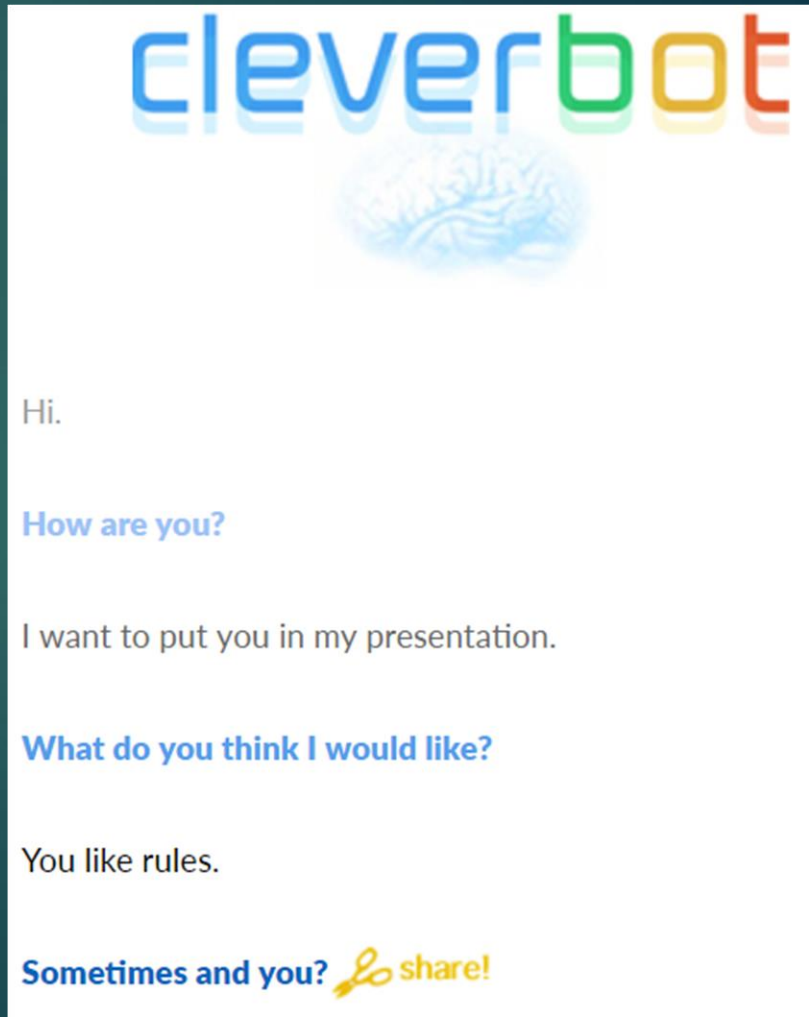
github.com/ricsinaruto/Seq2seqChatbots

Chatbot types

Neural

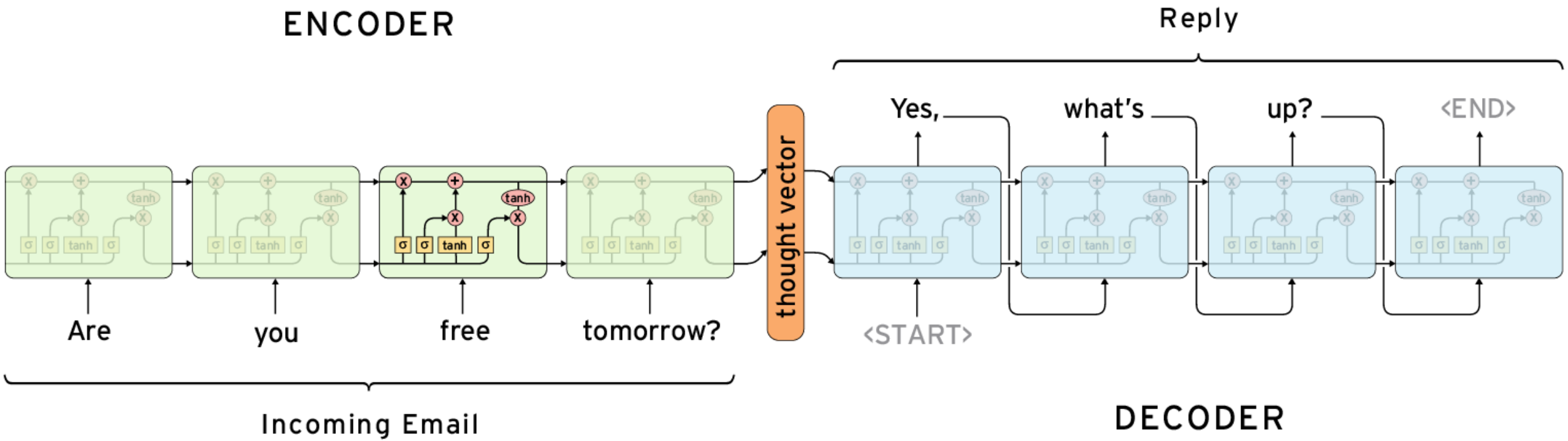


Rule-based



Seq2seq (2014)

$$\frac{1}{S}\sum_{T,S\in S}\log\left(p(T|S)\right)$$



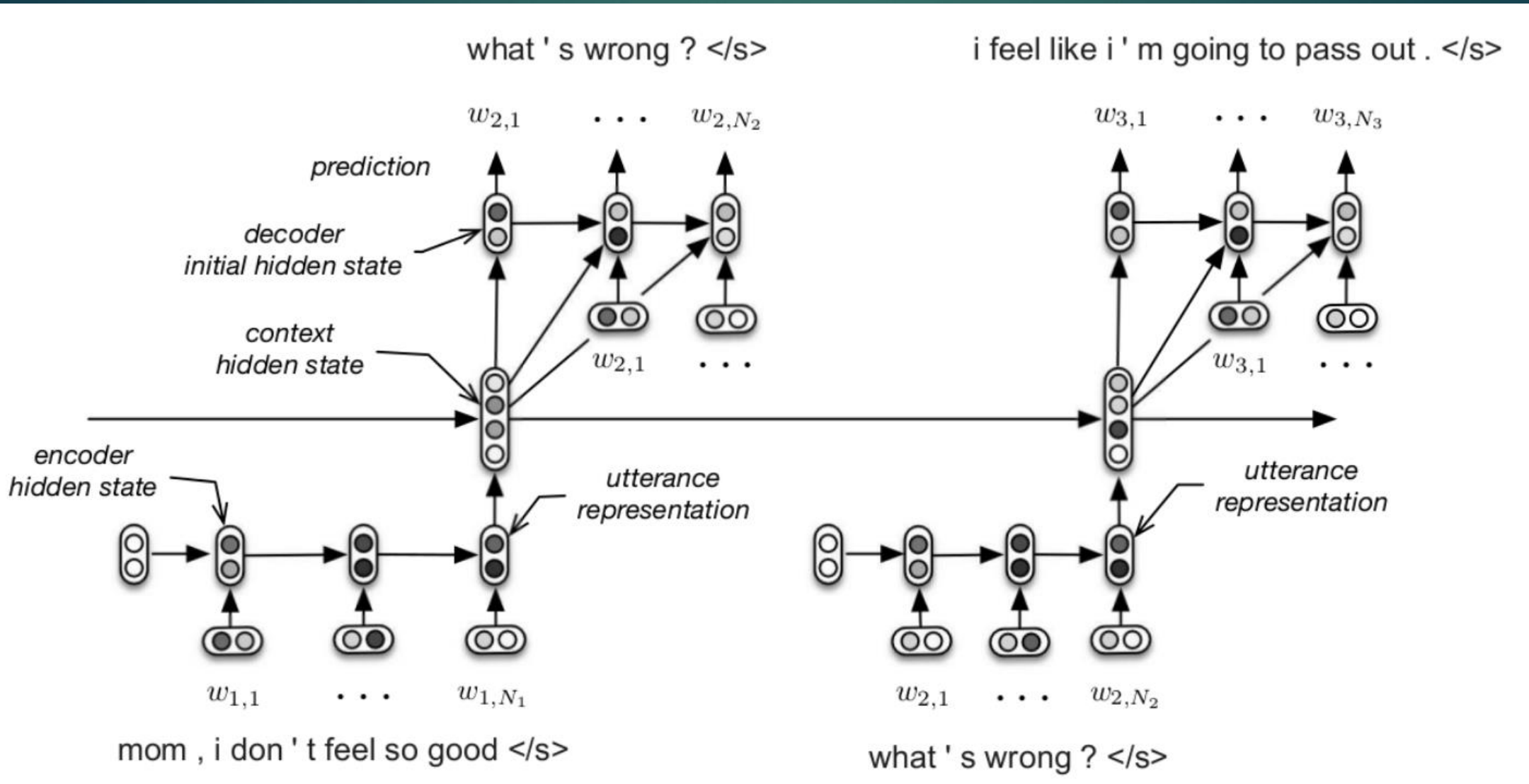
Datasets

- Twitter, Reddit
- Movie subtitles
- Crowdsourced conversations

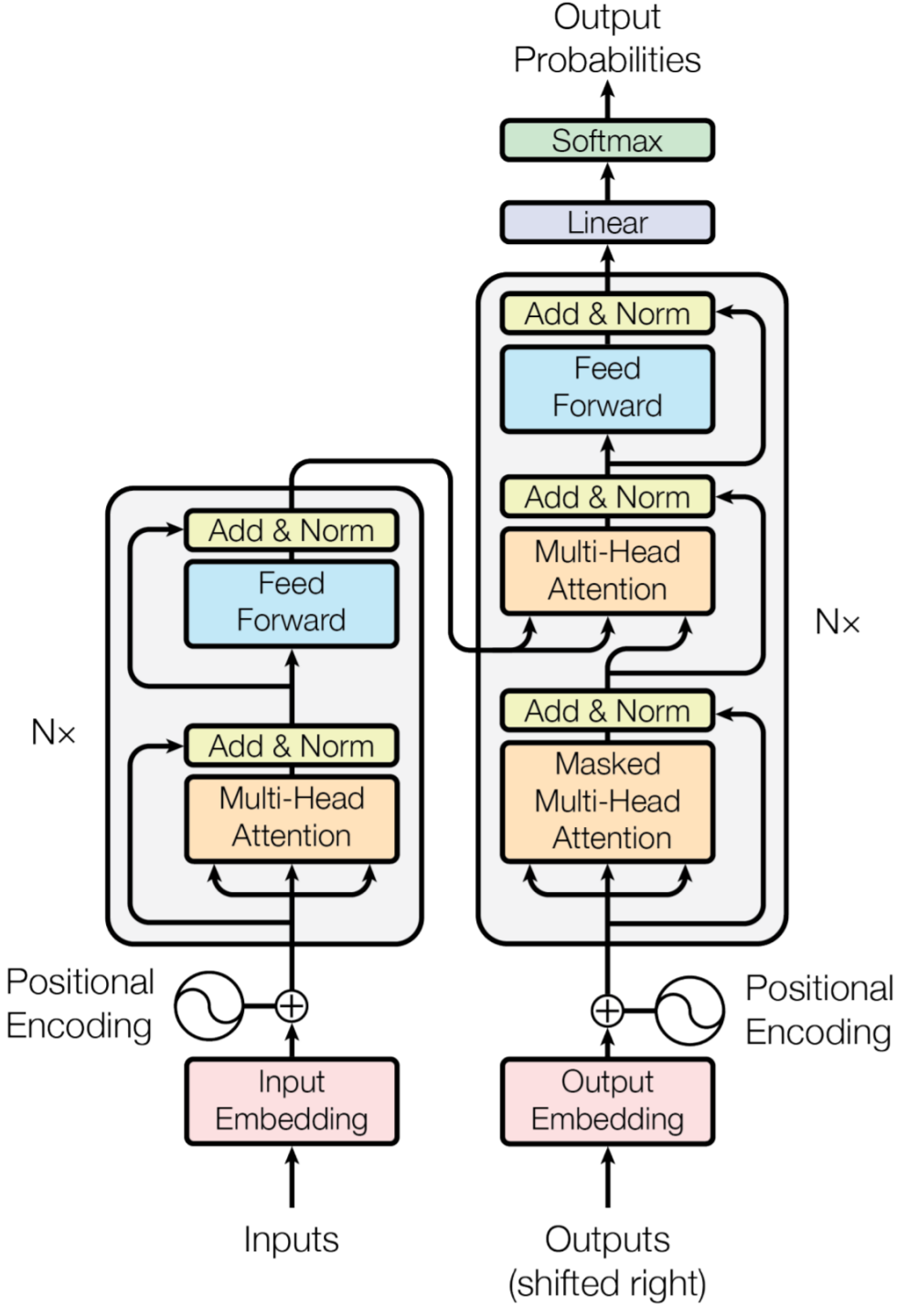
Evaluation

- Compare to target utterance (embedding, perplexity)
- Diversity (entropy)
- Human evaluation

HRED (2015) https://arxiv.org/pdf/1507.04808.pdf



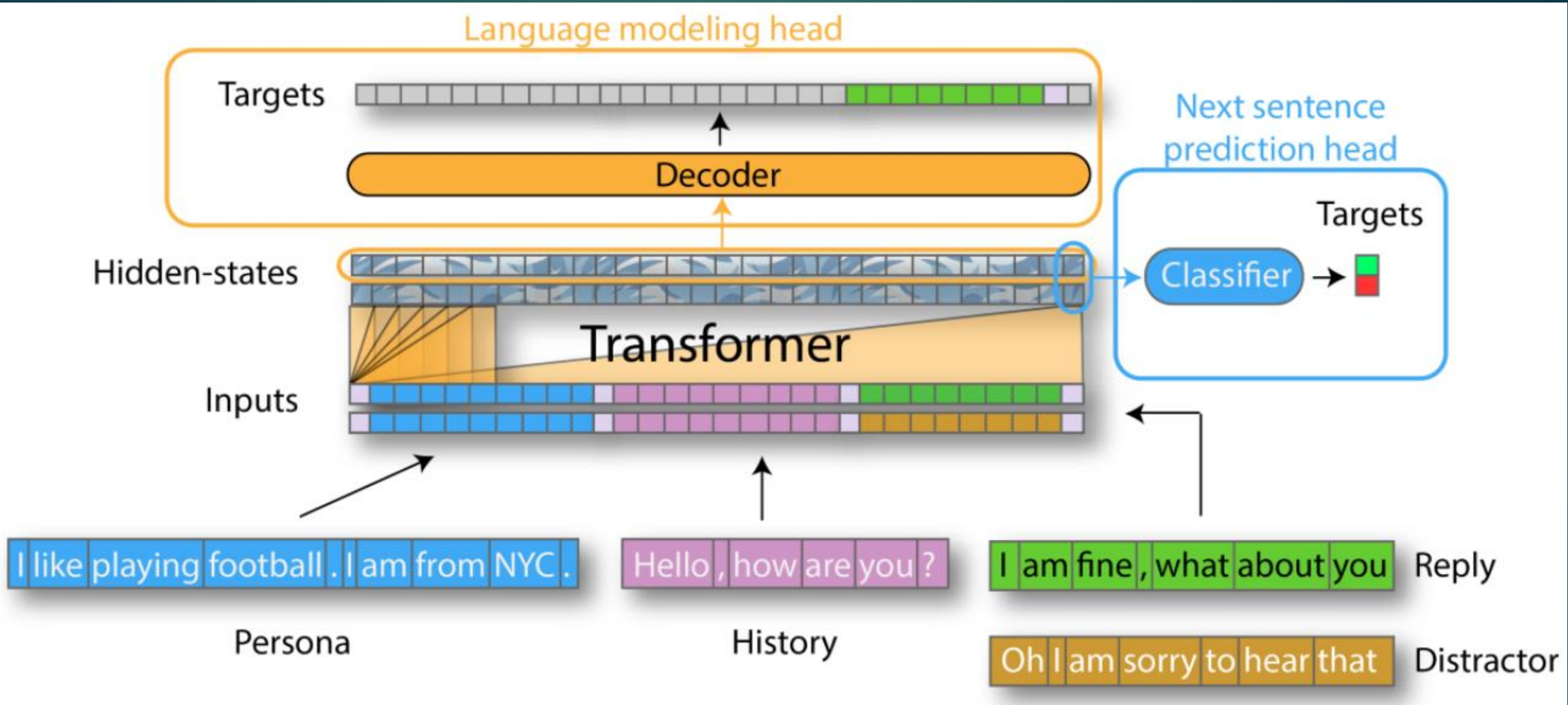
Transformer (2017)



http://jalammar.github.io/illustrated-transformer/

GPT-2 (2019)

https://medium.com/huggingface/how-to-build-a-state-of-the-art-conversational-ai-with-transfer-learning-2d818ac26313



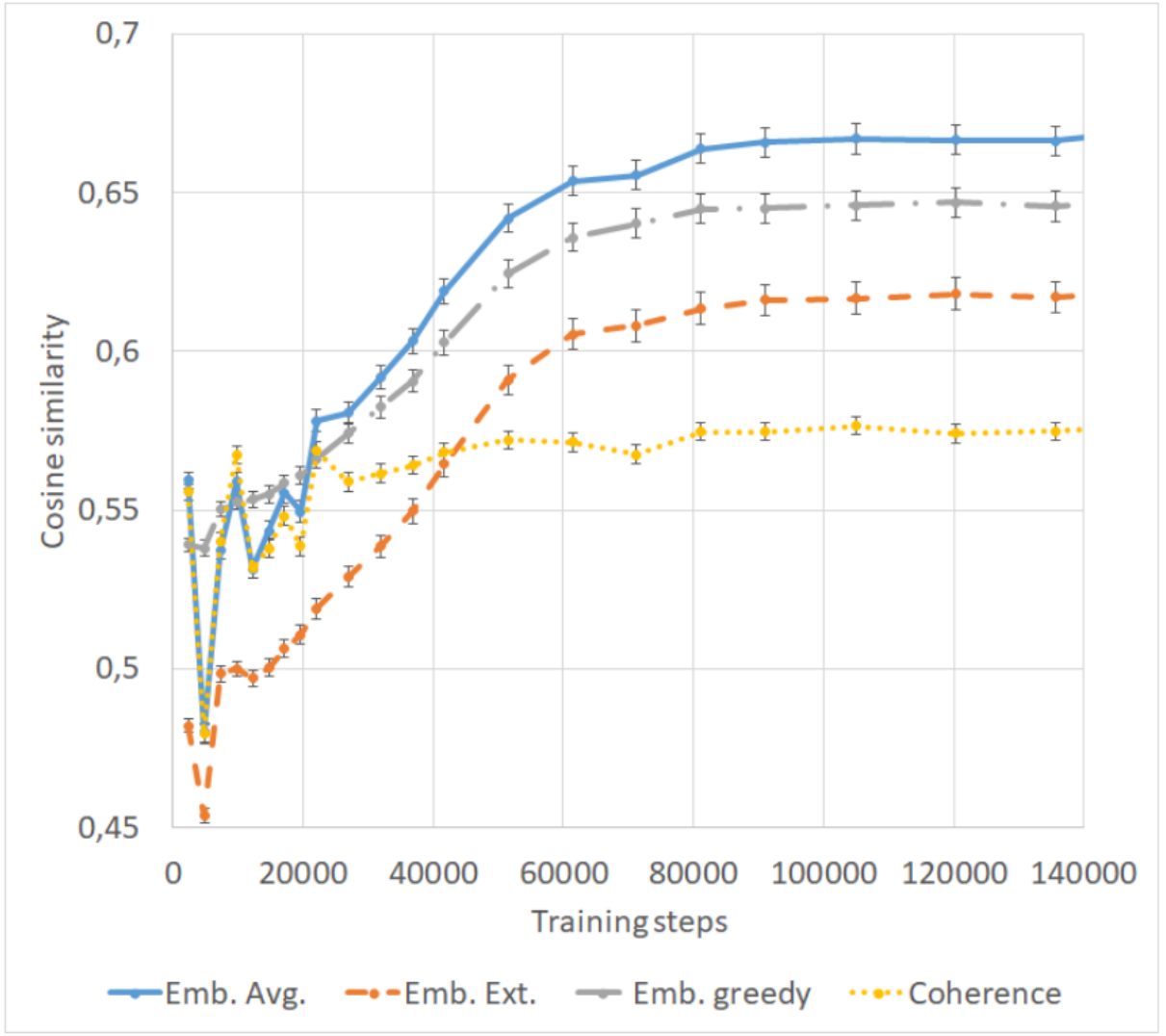
Issues: Data

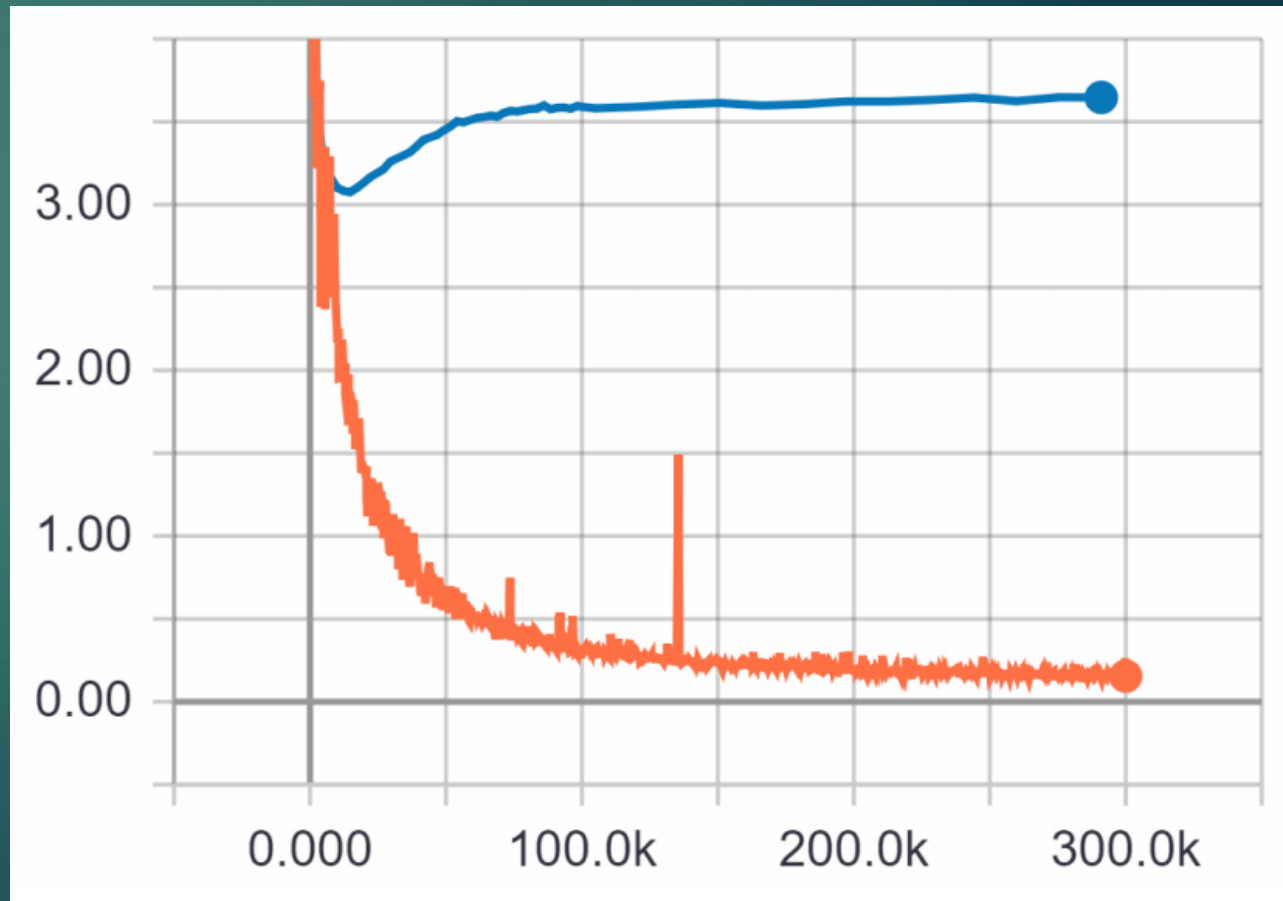
- Not natural
- Not enough

Issues: Loss function

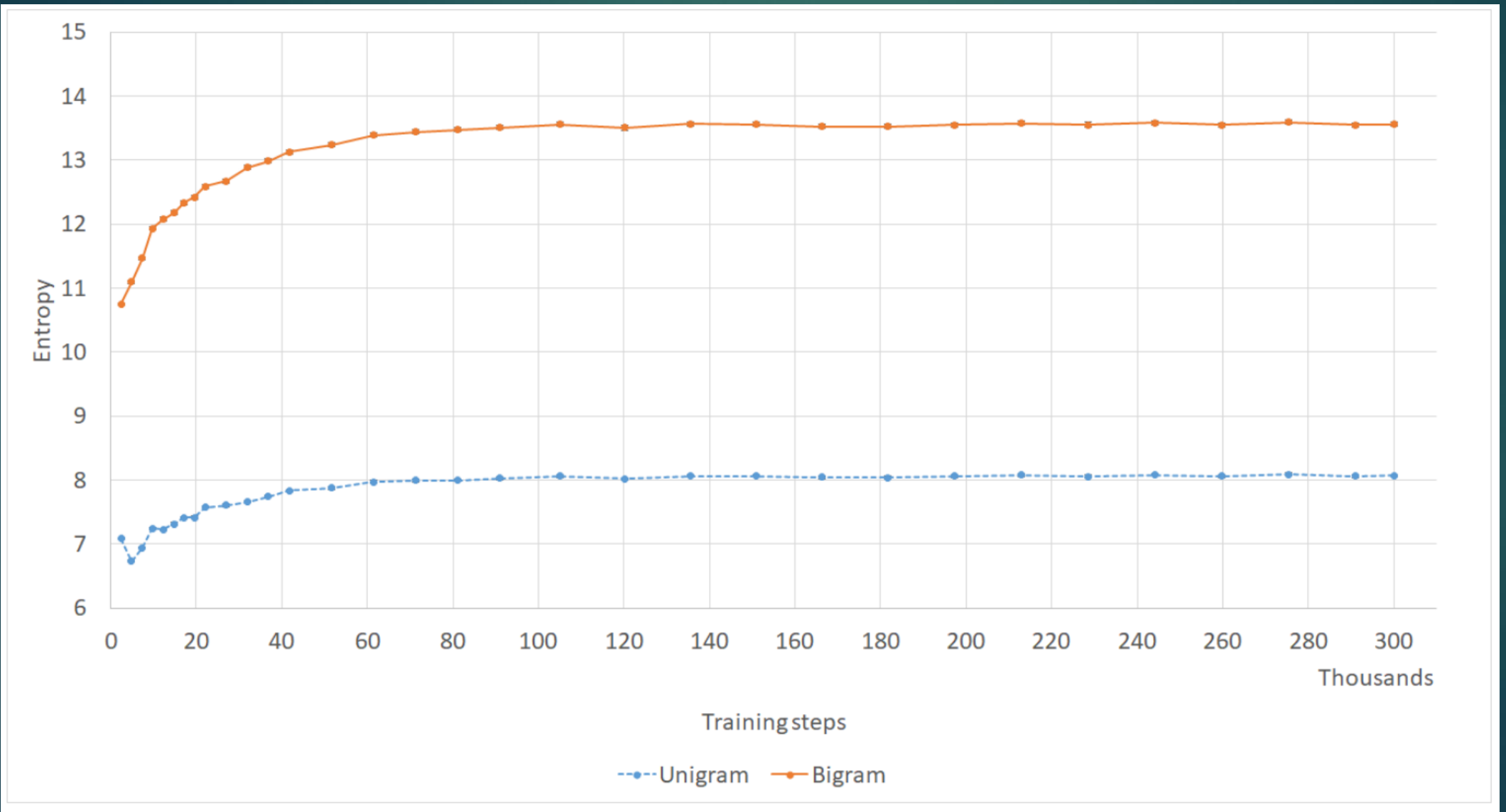
- Single-turn enc-dec architectures generate boring/generic responses
- One-to-many
  - How are you?
- Many-to-one
  - I don't know?
- Model doesn't perform best at validation loss minimum

Issues: Metrics





Issues: Metrics



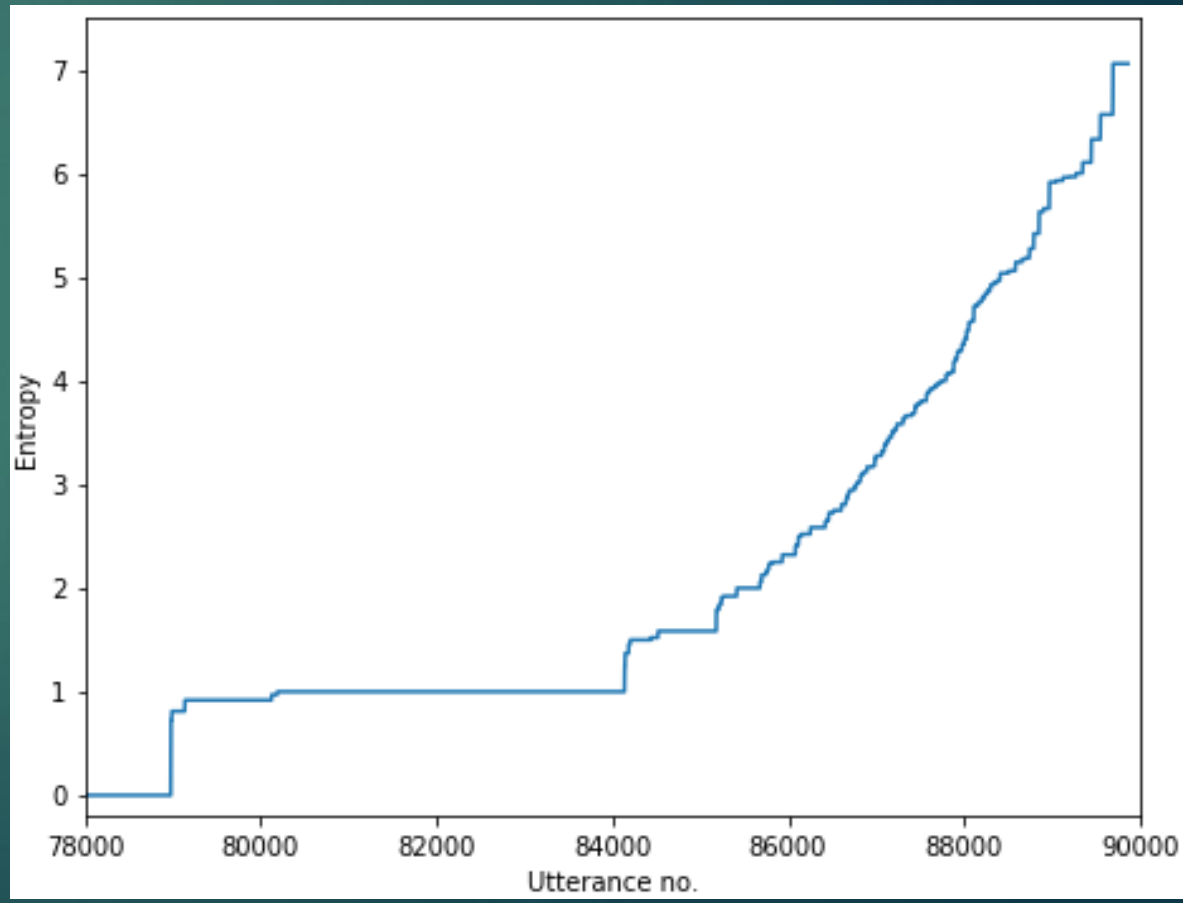
Ideas: Filtering based on entropy

https://arxiv.org/abs/1905.05471

- Remove problematic sentences from the dataset
  - One-to-many, many-to-one
- Entropy

$$H_{\text{src}}(t, D) = -\sum_{(s_i,t)\in D} p(s_i|t)\log_2 p(s_i|t)$$

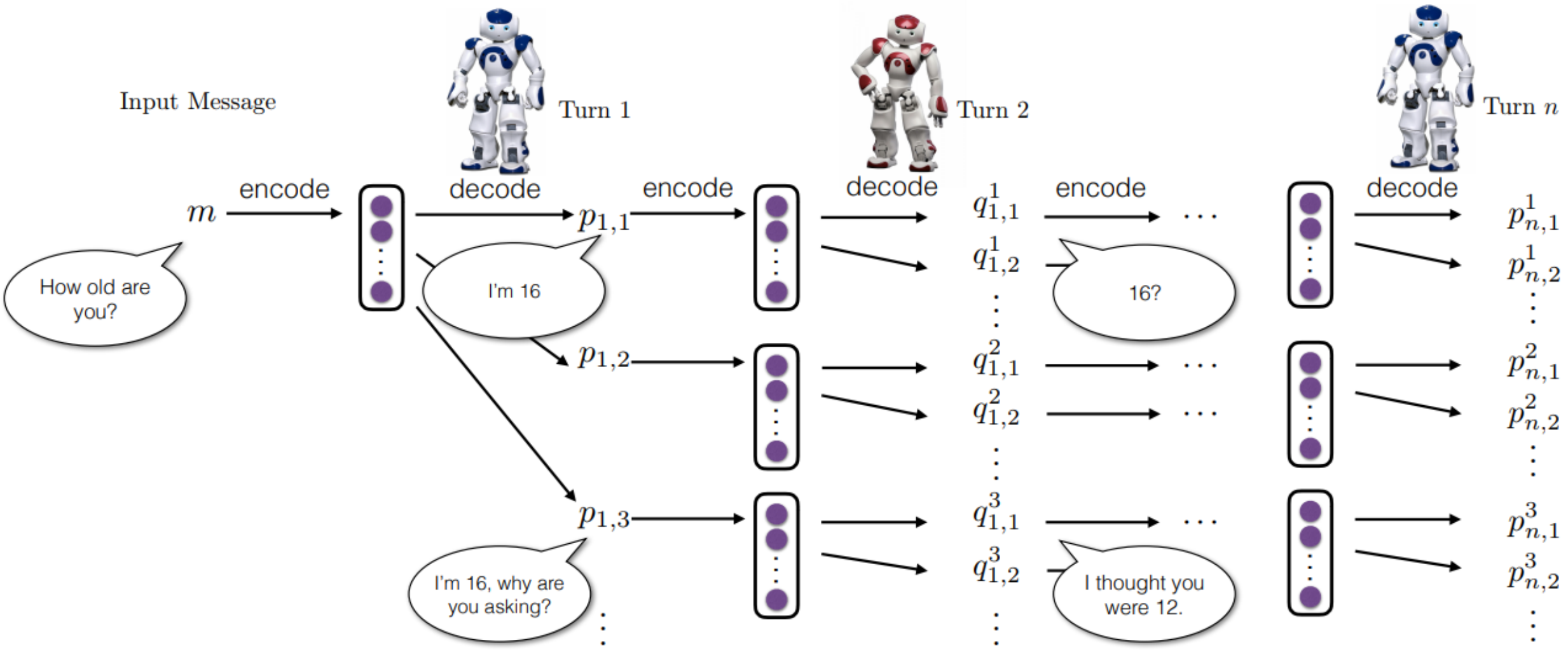
$$H_{\text{tgt}}(s, D) = -\sum_{(s,t_i)\in D} p(t_i|s)\log_2 p(t_i|s)$$



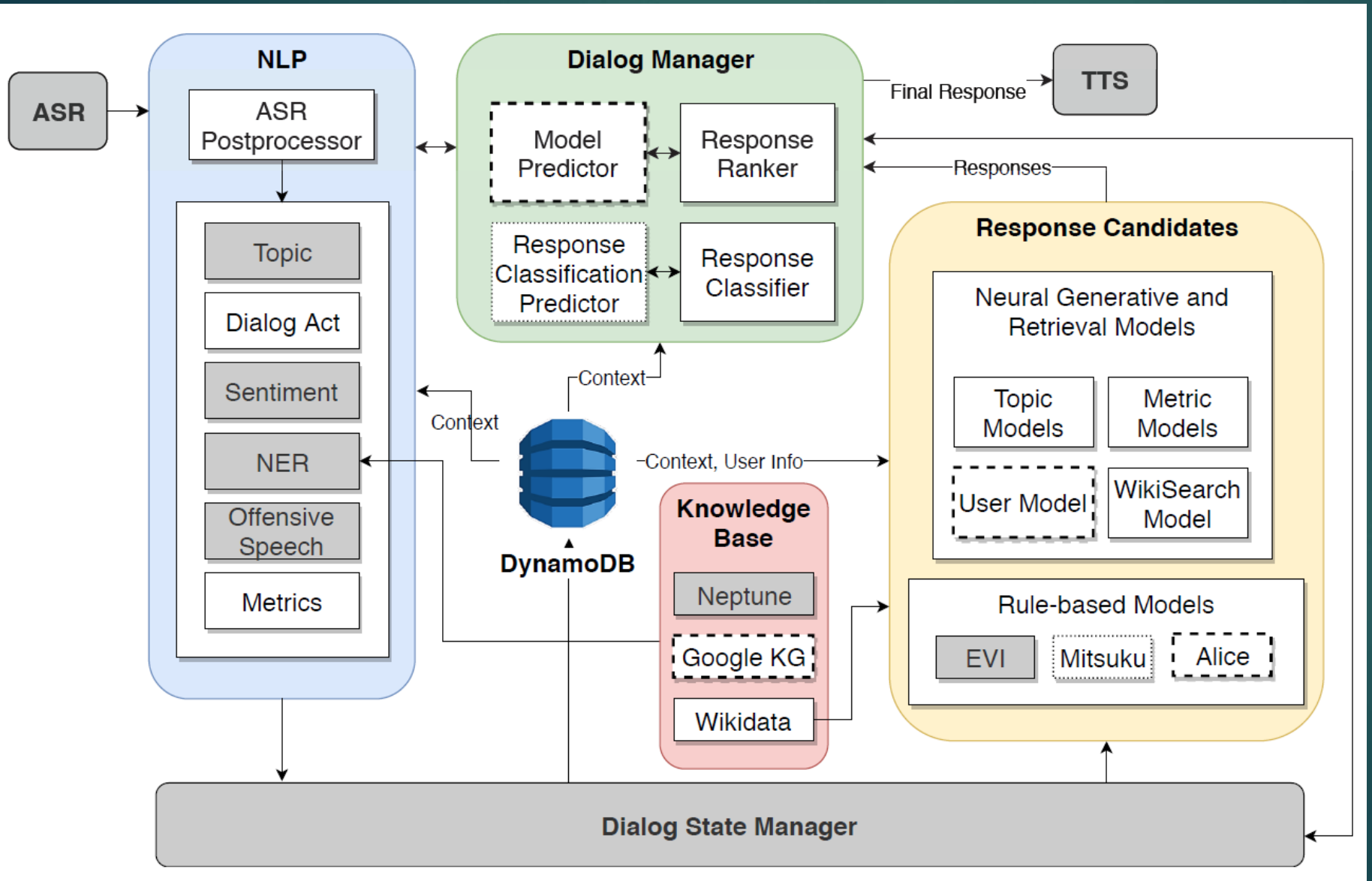
Ideas: Features

Ideas: self-play

https://arxiv.org/pdf/1606.01541.pdf



Amazon Alexa: A dialog system



Köszönöm a figyelmet!

- https://github.com/ricsinaruto/Seq2seqChatbots/tree/master/docs/meetup
  - Slide-ok és ParlAI tutorial