

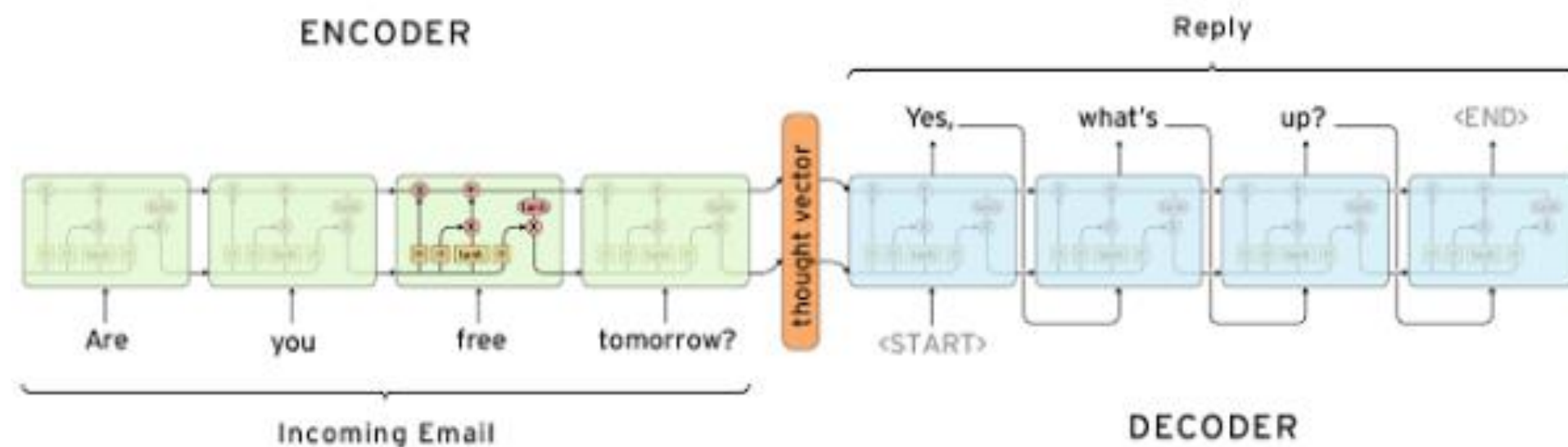
Chatbot: Hi, how are you?

PROBLEM

Chatbot:

- Capture an input sentence from users and perform text processing
- Apply and Benchmark Sequence to Sequence model variations to predict output sentence
- Best prediction is displayed via a chatbot web interface.

Approaches: Sequence to Sequence learning



- Neural Machine Translation by Jointly Learning to Align and Translate: <https://arxiv.org/abs/1409.0473>
- Evaluating Language Model: <https://courses.engr.illinois.edu/cs498jh/Slides/Lecture04.pdf>
- Sequence to Sequence Learning: <https://papers.nips.cc/paper/5346-sequence-to-sequence-learning-with-neural-networks.pdf>
- A Neural Conversational Model: <https://arxiv.org/abs/1506.05869>

DATASET

Data Sources:

- Subtitle dataset: Cornell's movie subtitles dataset with about 140,000 input-output sentence pairs.
- Twitter dataset: twitter conversation dataset with about 130,000 input-output sentence pairs.

Data Processing:

- Vocabulary Size: 8000 (Rare word replaced with "unk")
- Input Sentence Length: 20 (padding "<pad>" is used)
- Output Sentence Length: 20 (padding "<pad>" is used)
- End of Sentence "<eos>" is used

MODEL CONFIGURATIONS

Network Architecture:

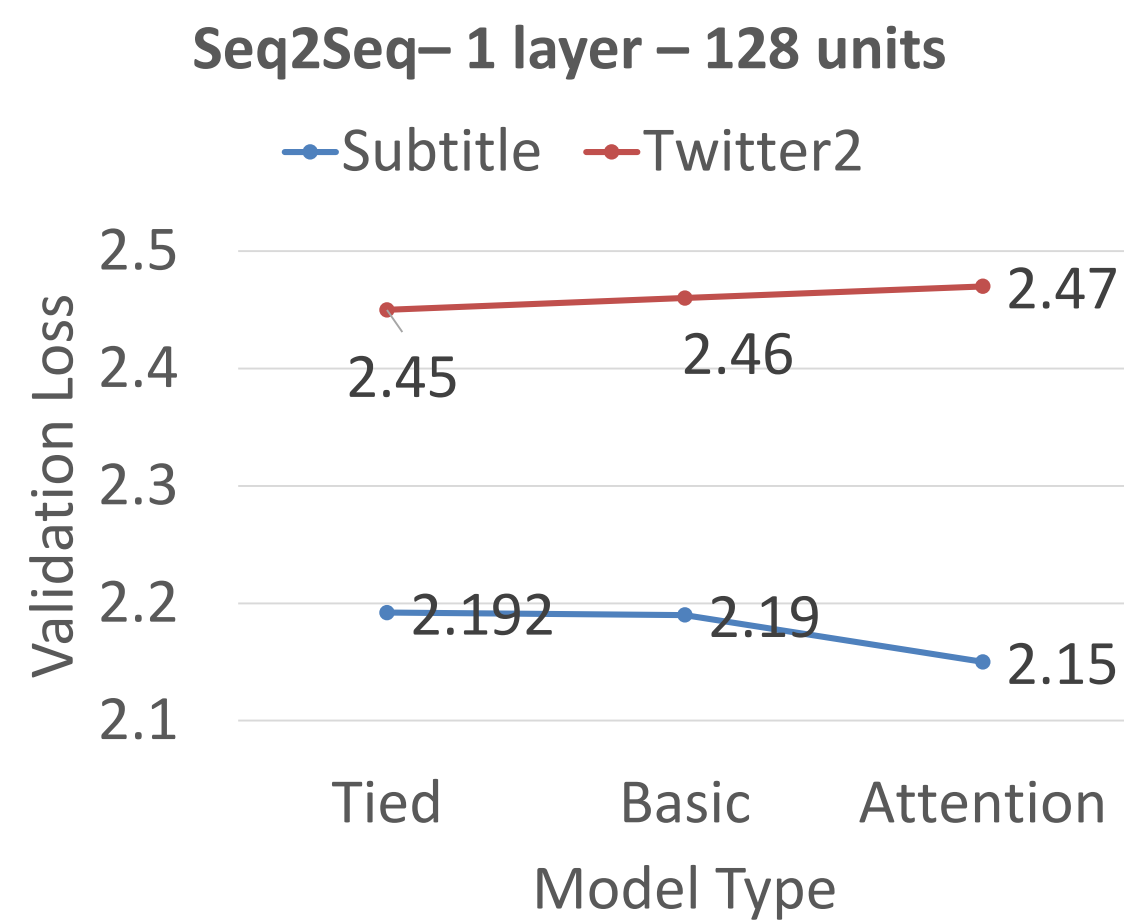
- Model Type: Basic, Tied, Attention Seq2Seq
- No. units in LSTM: 128, 256, 512, 1024
- No. stacked LSTM layers: 1, 2, 3

Network Training:

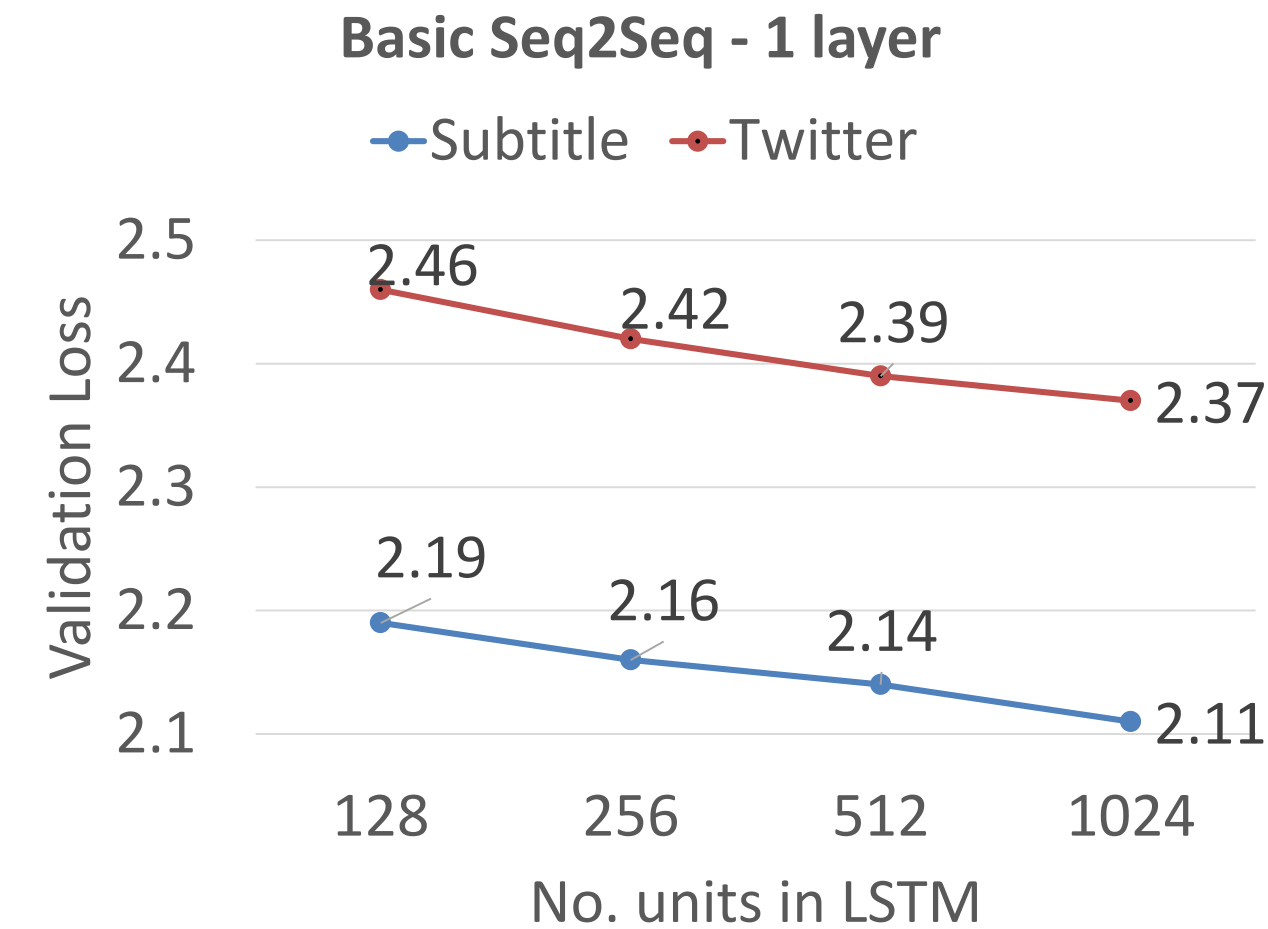
- Loss function: cross-entropy loss for sentence
- Optimizer: Adam
- Learning Rate: 0.001
- Batch Size: 64 -> 1,000+ batches per epoch
- No. epochs: 20

EXPERIMENT RESULTS

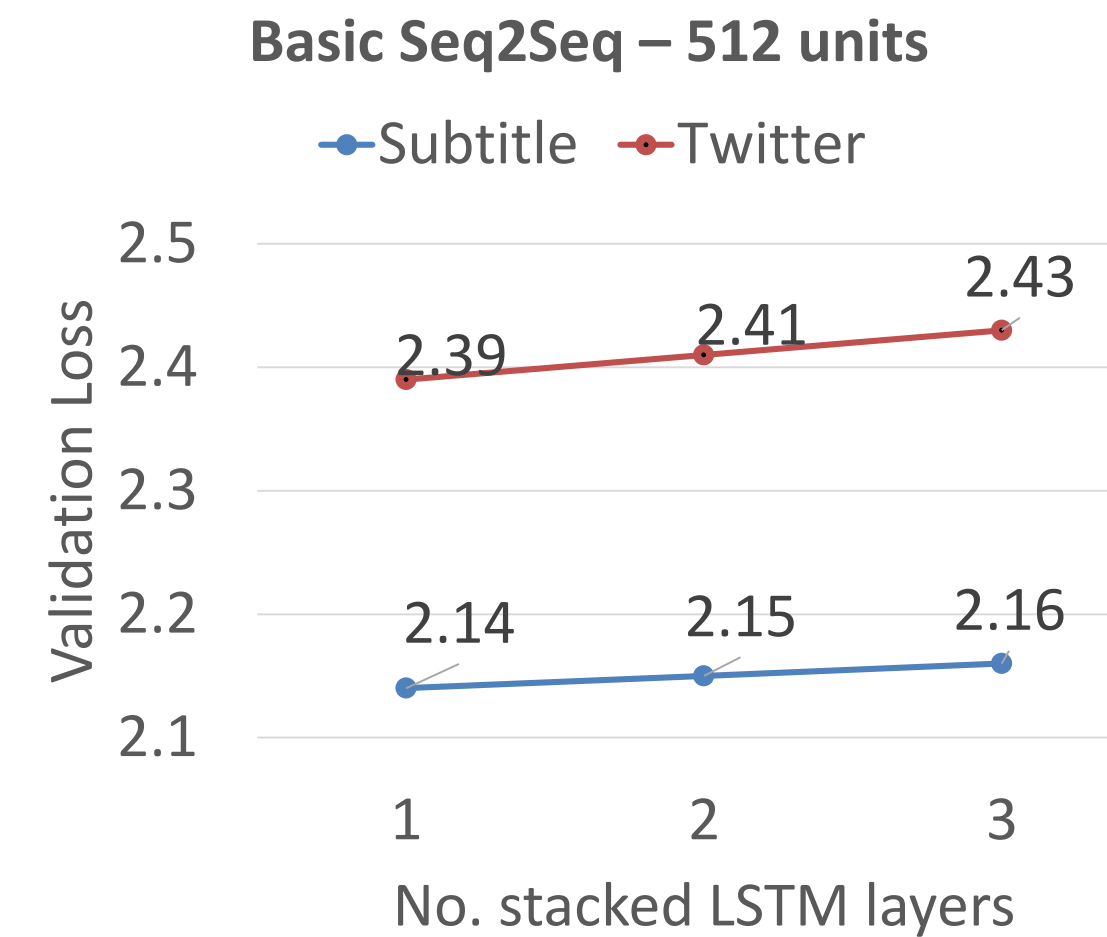
Model Type:



No. units in LSTM:



No. stacked LSTM layers:



OUTPUT EXAMPLES

Good:

- **Input:** happy birthday
- **True Output:** thank you
- **Predicted Output:** thank you

- **Input:** i love you so much
- **True Output:** i love you more
- **Predicted Output:** i love you too

Average

- **Input:** what do u like most about being part of this show?
- **True Output:** being part of the show
- **Predicted Output:** the best of the world

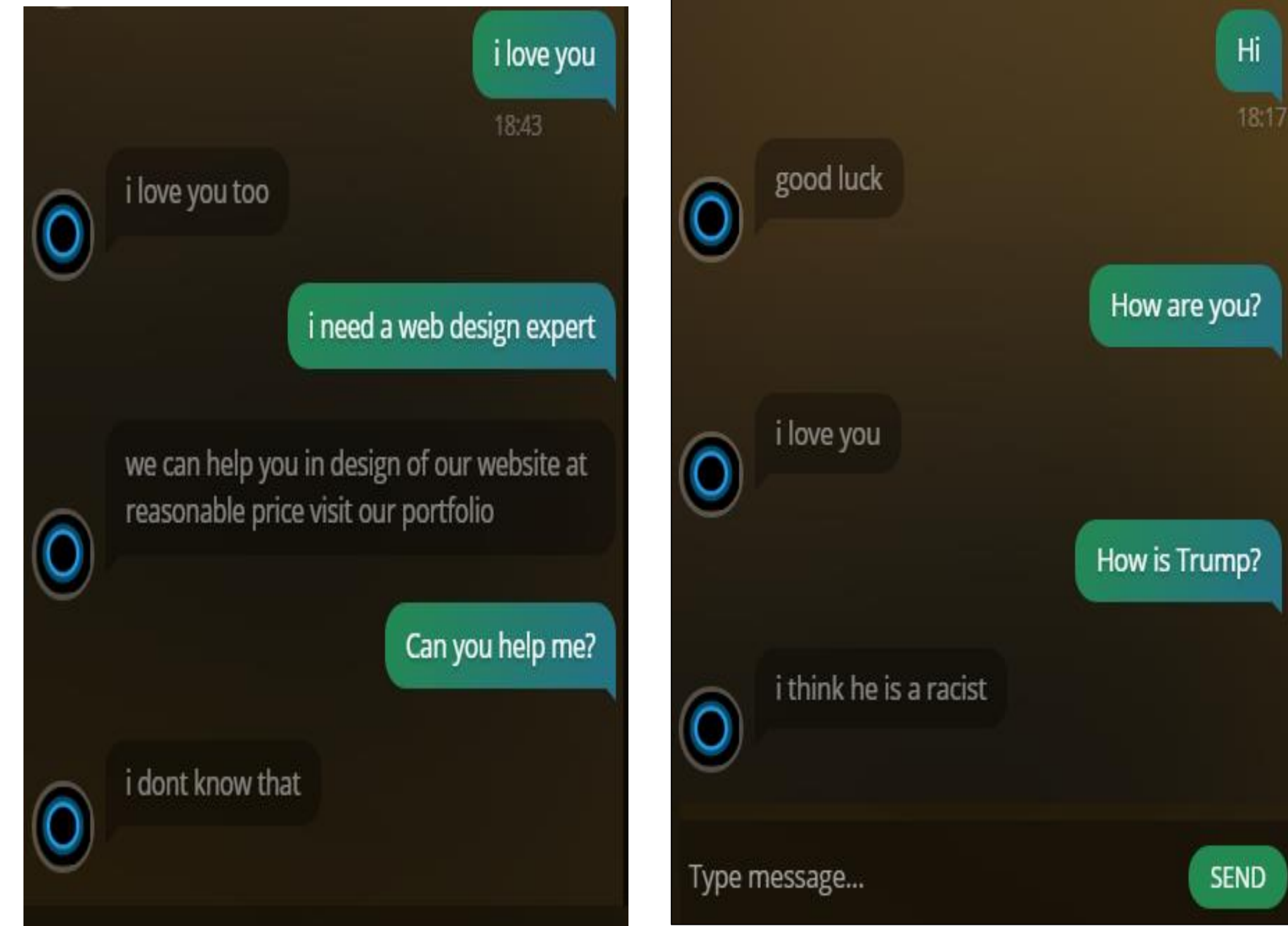
- **Input:** anyone know how to web design or know someone who can?
- **True Output:** we are experts in services in our portfolio email to leadscom
- **Predicted Output:** we can help you in design of our website in reasonable price visit our portfolio

Bad:

- **Input:** let me know if youre down for this
- **True Output:** very down lets do it
- **Predicted Output:** i think you are a good one

- **Input:** am i allowed to write bernie name on the ballot amp vote?
- **True Output:** in the name of democracy yes
- **Predicted Output:** you are a racist

DEMO



LIMITATIONS

- Due to computing resource constraints, we cannot train Seq2Seq model for larger dataset or larger network.
- The model currently does not have mechanism to maintain historical context (i.e. past sentences)

CONCLUSIONS

- Tied weight have little effect on the Seq2Seq chatbot model
- Experiment with Attention Mechanism is inconclusive
- Adding more units improve the model's learning capacity
- Adding more stacked layers does not improve the model, indicating overfitting.
- The model can learn grammar well, but its capacity to learn semantic meaning seems to be limited on small dataset.
- However, it does show some potential to learn the context. Perhaps, experiment on larger dataset could improve the model.