How do you feel?

Making chatbots emotional

Ethan Mines, Samira Shaikh

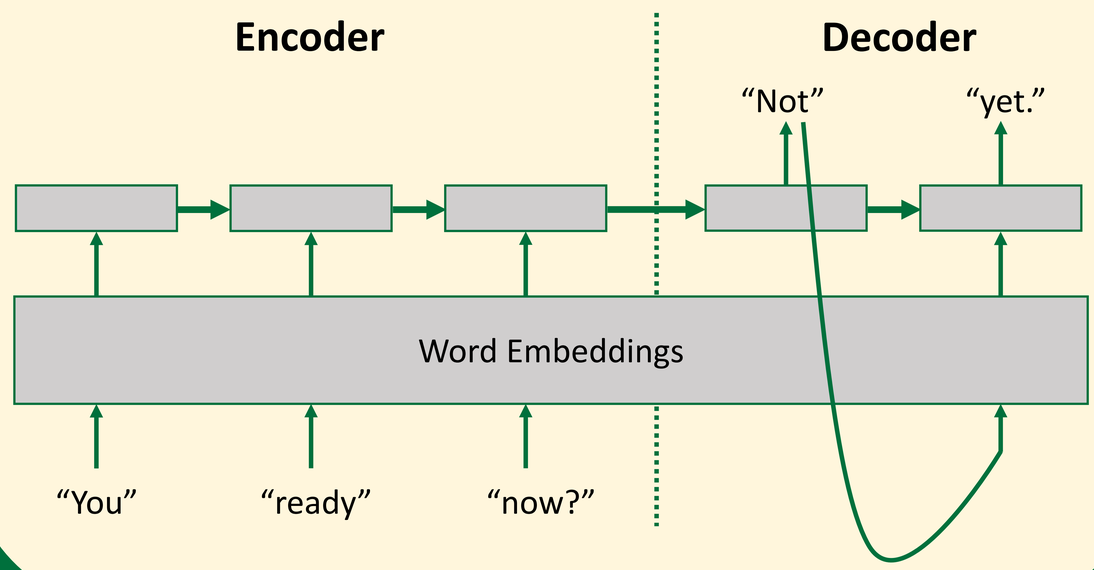
Visualization Center – The University of North Carolina at Charlotte

Chatbots

- Artificial intelligence models that converse or "speak"
- Applications
  - Personal Assistants
  - Customer service
  - Therapy
- Difficult to generate **emotional responses**, beyond simply "Yes," "No," "I don't know," etc.

Sequence-to-Sequence Neural Network

- Technique meant to emulate the neurons of the brain using a "network" of computations
  1. Receive a prompt
  2. Produce a response
  3. Compare response to correct one given by a human
  4. Update mathematical functions for more accurate output
- Encoder reads inputted prompt and "encodes" it in a vector
- Decoder produces response



Word Embeddings

- Means of representing words as mathematical vectors
- Similar words have similar vectors
- Semantic embeddings capture meanings
- Affective embeddings capture emotions

"Mom" → $\langle -0.54, 0.358, 1.001, \ldots \rangle$

"Mother" → $\langle -0.51, 0.363, 0.999, \ldots \rangle$

Methodology

Using a dictionary of approximately 14,000 words with assigned emotional values of Valence, Arousal, and Dominance (VAD), we test three models each with different embeddings. The first has a word's VAD values appended to its semantic embedding (**VAD Append**) [1]. The other two have affective content "injected" into semantic embeddings, one via counterfitting (**Counterfitted)** and the other via retrofitting (**Retrofitting**) [2].

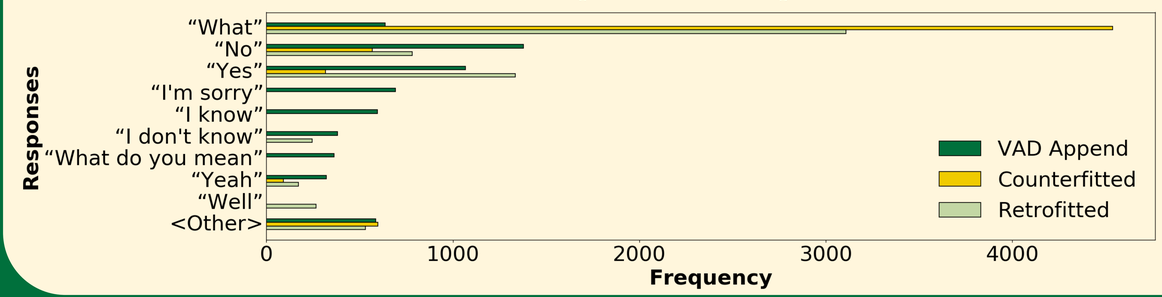
We use conversations drawn from movie scripts for training data. The model is tested using a separate set of approximately 6,000 prompts.

Discussion

All three models produce mundane responses to the test prompts, such as "What?" and "I'm sorry." This indicates that affective embeddings, in themselves, are insufficient for generating emotional responses. Future experiments will utilize techniques beyond affective embeddings. For example, during training, the model can be scored for emotionalness in addition to accuracy.

The **VAD Append** model did generate somewhat more diverse responses than the others, likely due to its larger embeddings (1027 dimensions vs. 303).

Common Responses by Model



Sample Responses

Prompt	VAD Append	Counterfitted	Retrofitted
"I'm going to be in this actor's workshop, and I'm hoping to start going on auditions soon. I'm so excited to finally have some free time. We have to get together this summer!"	"What are you going to do?"	"What?"	"I won't."
"But what if Andy gets another dinosaur? A mean one? I just don't think I can take that kind of rejection."	"I don't think so."	"What do you think?"	"I don't know."

Acknowledgements

This research was made possible by a National Science Foundation grant and the high-performance computing resources provided by the University Research Computing department at UNC Charlotte.

References

[1] Nabiha Asghar, Pascal Poupart, Jesse Hoey, Xin Jiang,and Lili Mou. 2017. Affective neural response generation. *CoRR*, abs/1709.03968.

[2] Sopan Khosla, Niyati Chhaya, and Kushal Chawla. 2018. Aff2vec: Affect–enriched distributional word representations. *arXiv preprint arXiv:1805.07966*.