Comments Without Food

Generating Fake Restaurant Reviews Using RNN

Tianxiao Hu, Yi-Fang Lee, Zhenbang Chen, Wenyuan Gao
{tianxiao.hu, yifang_lee, zhch, wenyuan_gao}@berkeley.edu

Introduction

Recurrent neural networks (RNN) can be used to learn sequential data and generate similar data. In this project, we focused on restaurant review generation using RNN. The model learns from restaurant reviews on *Yelp.com* (which is wrote by human). Our goal is to generate reviews realistic enough to fool native speakers.

We follow two distinct branches in pursuing this goal: character-level language model and word-level language model. For character-level models: we take vanilla two-layer stacked LSTM network as our baseline. Then we improved the network structure with additional embedding layer and attention layer. For our word-level model, it starts with pre-trained word2vec embedding layer, along with two stacked GRU layers. Also, dropout is used to prevent overfitting. Finally, we compared our model with the state-of-the-art GPT-2 language generation model (simplified version).

We used the *Yelp Dataset* to generate fake reviews. The dataset contains 3+ million 1-5 star restaurant reviews. For each network architecture, we train it on two datasets: 1-star reviews and 5-star reviews for better comparison.

Model Overview

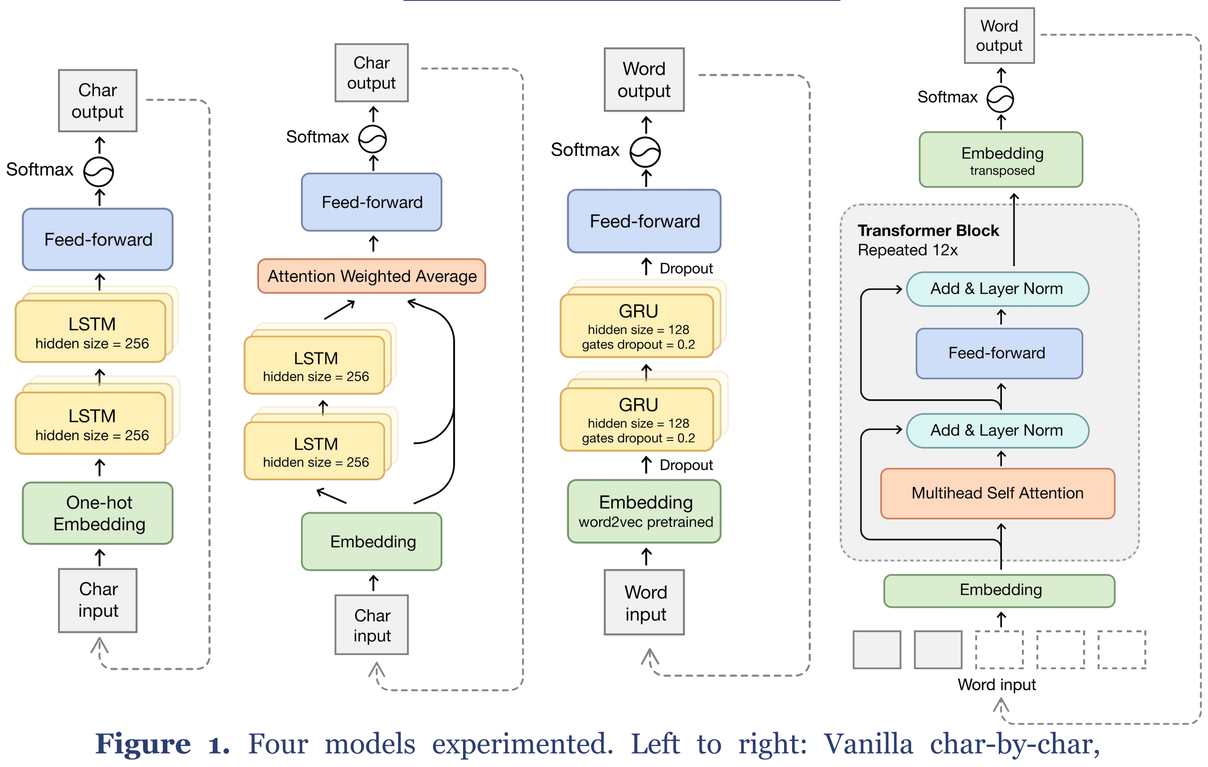

Figure 1. Four models experimented. Left to right: Vanilla char-by-char, Char-by-char with attention, Word-by-word, and GPT-2 (simplified).

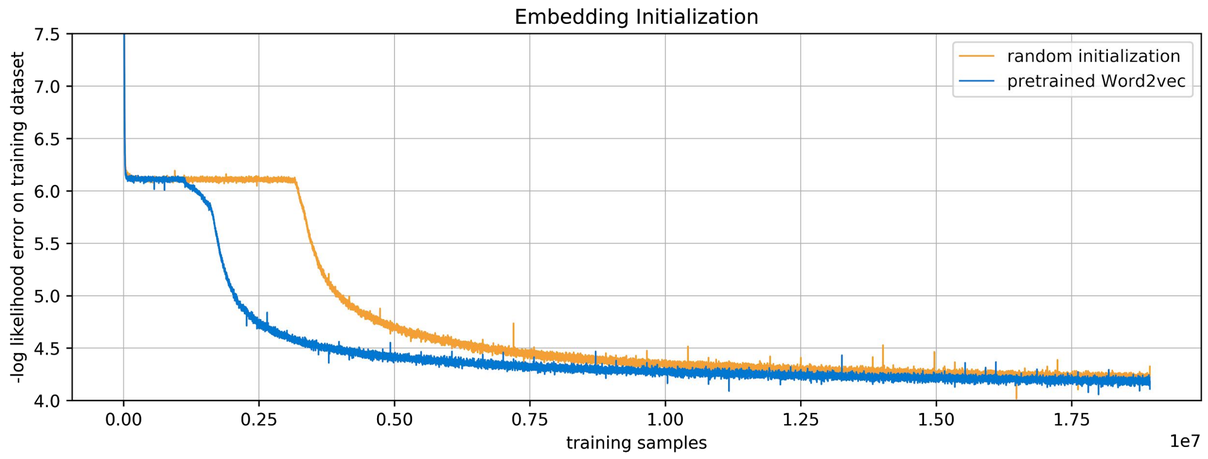

Figure 2. Comparison of training with versus without pre-trained Word2Vec embeddings

As displayed in Fig 1, the basic char-by-char model has two LSTM layers with 256 cells each followed by a feed-forward layer. The improved one is implemented with extra embedding and attention layer. The word-by-word model employs word embedding and GRU layers with dropout. The simplified GPT-2 model has 12 repeated transformer blocks with multihead attention structure.

As shown in Fig 2, we noticed that using pre-trained Word2vec embedding significantly improved training speed, and the performance was slightly better, compared to training tabula rasa.

Results and Conclusions

Five-star Reviews Generated	
Vanilla char-by-char	This is the best in Phoenix area! The service was amazing. It was nice and tasty and creative in a large group of friendly and attentive staff and no other omelet over the exception. I also order the potato, chicken and the chicken parmbialese.
Char-by-char with attention	I love this place. I have been going to this location and was very convenient. The food was amazing and the service is always good. The owner is super friendly and helpful. The service was fast.
Word-by-word	It was amazing. I had to try the el balls, and the seasoning was delicious. The amount of meat which was amazing and the sauce was the best, I have ever had. I will definitely be back. Biloxi fried ice cream and the fried banana. Both are delicious.
GPT-2 (simplified)	We like to think that this is the best Thai restaurant in the area and possibly in Mississauga as far as Thai is concerned. The food is certainly well prepared and the staff is extremely friendly. The ambience is calming and inviting. A traveler in need of a table? This is a must try!

One-star Reviews Generated	
Vanilla char-by-char	Will not come here again. We were disappointed that I was the only ones that the hostess stopped here. Because of the Kateron that there was a bad service and the food was bland and overcooked.
Char-by-char with attention	The food was okay. The rice was ok, but the waitress was pretty stupid for the first time. We were told they were out of bread. The sandwich was so tough they didn't have a couple of salads.
Word-by-word	If I could give this place 0 stars I would. I ordered a pepperoni salad and was told it was a small portion of meat. I was ordering it. I said it was supposed to be cooked. Also, I got a mouthful of chicken on the side and he was over cooked.
GPT-2 (simplified)	Worst dim sum place that I have ever been. They have black hair in their food and urine. The hostess had the whole restaurant and all of its personnel in front of her at the same time. It was incredibly rude and graphiced everything we did and where we were served.

Table 1. Comparison of generated sample texts

We use 20, 000 reviews for character-level language model training, and 100, 000 reviews for word-level. Each model was trained for ~25 epochs. It's noteworthy that we only fine-tuned the GPT-2 model using 1,000 reviews, which is enough to generate realistic comments.

Our results demonstrate that even two-layer RNNs are a fully feasible way of generating grammatically correct, convincible fake restaurant reviews.

References

[1] Yelp.com. (2019). Yelp Dataset. [online] Available at: https://www.yelp.com/dataset [Accessed 17 Feb. 2019].
[2] Radford, Alec, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. "Language models are unsupervised multitask learners." (2019).
[3] Karpathy, Andrej. "The unreasonable effectiveness of recurrent neural networks." *Andrej Karpathy blog* 21 (2015).
[4] Yao, Yuanshun, et al. "Automated crowdturfing attacks and defenses in online review systems." 2017 ACM SIGSAC Conference on Computer and Communications Security. ACM, 2017.
[5] Mikolov, Tomas, Ilya Sutskever, Kai Chen, Greg S. Corrado, and Jeff Dean. "Distributed representations of words and phrases and their compositionality." In Advances in neural information processing systems, pp. 3111-3119. 2013.
[6] Radford, Alec, Karthik Narasimhan, Tim Salimans, and Ilya Sutskever. "Improving language understanding by generative pre-training." (2018).

Berkeley
UNIVERSITY OF CALIFORNIA