

Multi-Task Learning for Speaker-Role Adaptation in Neural Conversation Models

Yi Luany Chris Brockettz Bill Dolanz Jianfeng Gaoz Michel Galleyz
Department of Electrical Engineering, University of Washington
Microsoft Research

Motivation

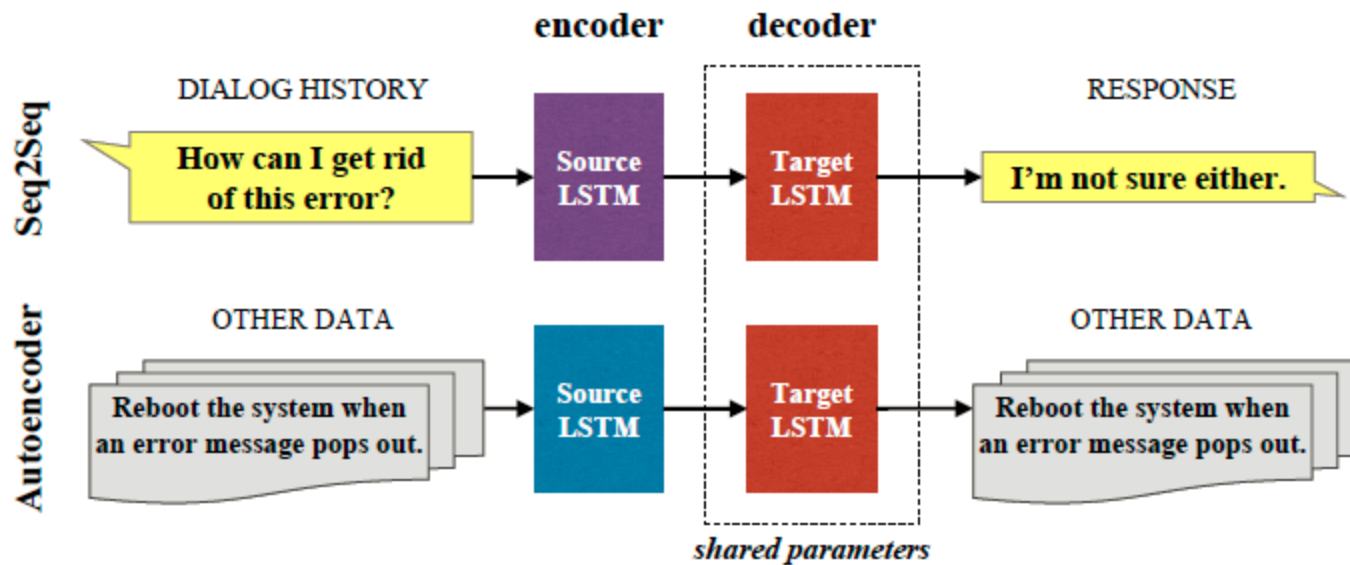
Multi-Task learning can make the agent exhibit identifiable role-specific characteristics by incorporating non-conversational data.

Model

- MTask-S: N models for N users (be inefficient in both memory and computational cost)
- MTask-M: introduce user embeddings to Seq2Seq model

User embedding is initialized randomly during training and is updated by auto-encoder training together with the decoder LSTM parameters.

Model Framework



Training Process

Training procedure of Multi-task learning:

1. Randomly initialize SEQ2SEQ and AUTOENCODER encoder parameters.
2. Train SEQ2SEQ model until dev set performance converges in perplexity.
3. **While** not dev set performance converged in perplexity **do**:
 - (a) Randomly pick a batch of samples from general conversational data.
 - (b) Compute loss and gradient for SEQ2SEQ task and update parameters.
 - (c) Randomly pick a batch of samples from non-conversational data of the target user.
 - (d) Compute loss and gradient for AUTOENCODER task and update parameters.
4. Choose the best model based on SEQ2SEQ perplexity performance on dev set.

Conclusion

User embedding can make the dialog system exhibit the role-specific characteristics better. And multi-task with auto-encoder can learn the user embedding better with utilize unlabelled data.