The paper I chose for this assignment is: **A Neural Conversational Model by Oriol Vinyals and Quoc V. Le**. It was published in the proceedings of **ICML 2015** and can be found on arXiv here: https://arxiv.org/pdf/1506.05869v2.pdf. It has been **cited by 234** publications till now according to google scholar. It is based on seq2seq learning with neural networks (by Sutskever, Vinyals and Le (2014), discussed in class few weeks ago).

## Summary

The paper talks about an end-to-end framework based on using neural networks for mapping input sequences (statements) to output sequences (responses). This approach differs from conventional conversational systems because it lacks domain knowledge, can be trained end-to-end and requires fewer hand-crafted rules. The model predicts the next sentence given the previous sentence / sentences in a conversation.

A sequence to sequence, called seq2seq framework which takes input of one token at one timestamp and outputs one token at one timestamp. The framework is established using RNNs. The training phase involves learning the given true output through back propagation. In language tasks the objective is to decrease the perplexity and in this task as well the model is trained to maximize the cross entropy of the correct sequence using the context.

## Datasets Used

1) IT Helpdesk Troubleshooting dataset. (Training set: 30M tokens, validation set: 3M tokens)
2) OpenSubtitles Dataset (Tiedermann, 2009). (Training set: 923M tokens, validation set: 395M tokens)

## Advantages

- Since we are using neural networks, it is easy to use the same model for machine translation, question/answering and conversations without major changes in architecture.
- Can generate simple and basic conversations using very few hand-crafted rules unlike conventional bots.
- Able to extract knowledge from domain specific dataset (IT Helpdesk, finding solution to a technical problem) and from a large, noisy and general domain dataset of movie subtitles (common sense reasoning).

## Disadvantages

- Unlike easier tasks like Translation, this model does not completely solve the problem of modelling dialogue because:
    - The objective function being optimized does not capture the actual objective of human communication.
    - Since it is a purely unsupervised model, it is unable to ensure consistency and general world knowledge.
- Lack of personality. Hence cannot pass the Turing test. A paper by Microsoft talks about this limitation and introduces persona-based neural conversational models. (Find it here: https://arxiv.org/abs/1603.06155)
- No performance measure to calculate quality of models except manual inspection and perplexity.

## New Research Ideas

In the light of the recent 'ICLR 2018 Reproducibility Challenge' by Joelle Pineau of McGill University. I would like to reproduce the results of this paper myself. (example, http://neuralconvo.huggingface.co/)