# The SmartChatBot

## Your Friend who understands any document and answers your questions.



• **Introduction and Motivation**: Sometimes we get tired of listening to blunt answers from a chatbot. It is here that the Smart bot comes to our rescue. I want to build a network which can answer about any given text, a web page, an article or even a book, anything you give it. Suppose one of our friends has already read a long article and has understood that, and now we want to start reading the same article, taking help from your friend and asking him to summarize the article would really help us to get started on the document and also get a feel of the document before you actually decide to read it. How about a question answering machine on Kindle books? Wouldn't it be wonderful when someone can answer any question you have on the book? This kind of expert system will boost our productivity and aid us immensely to master any topic. These are some of my motivations to build this network.

• **Literature Review**: Prior studies have identified common models for question answering using memory networks (Weston et al. 2015 )[4]. They also introduce the concept of Dynamic knowledge base in the form of LSTM networks. Work has been done to improve this model by introducing an attention mechanism and training it end-to-end. (Sukhbaatar et al. 2015)[2]. Evaluating Reading comprehension and providing answers has been studied in progress towards AI Complete. (Westin et al. 2015)[3]. Previous memory networks have been improved by extracting global and salient features. (Yue et al. 2017)[7]

• **Deliverable**: I would design one hypothesis which will help me predict the answers to simple questions like the following:

- *Example 1*

Mary gave the cake to Fred.

Fred gave the cake to Bill.

Jeff was given the milk by Bill.

Who gave the cake to Fred? A: Mary

who did Fred give the cake to? A: Bill

- *Example 2*

Sandra travelled to the office.

Fred is no longer in the office.

Is Fred in the office? **A: no**

Is Sandra in the office? **A: yes**

My hypothesis would be based on Memory Networks, because these help in remembering the content of the document, like how the human do while reading an article. I would like to design a network whose answers are reasonable to humans and help them understand the given text.

• **Methodology**: I would like to design two networks based on LSTMs and Augmented Memory neural networks. LSTMs are already state-of-the-art in chatbot and Q/A systems, so this part would be mostly on implementation. The design of memory networks will focus on allowing a neural network to access an external data structure as the memory storage. It will learn where to retrieve the required memory from the memory bank in a supervised way. I would initially compare and contrast their performance. I finally would like to design ensemble from both the hypothesis and give the best prediction. I would also go through some research on going in Attention models and dynamic memory networks which are also a good class of models in Question Answering Systems.

• **Resources and Learning**: I need to collect good datasets. For initial working, I have a good dataset from facebook[1], but for eventual tasks I need to search/collect good datasets. I need to learn about Memory networks and LSTMs and their mathematical validation. Deep neural networks have always been a mystery to express mathematically, currently I am working on another project on Interpretable machine learning models, I would like to inculcate this in expressing the current problem mathematically. I am optimistic about this interpretability problem, but I am not sure whether I can complete it within the project span. I will definitely strive hard to my extreme.

• **Progress schedule (Milestones):**

1st Checkpoint: Oct 15 for learning and collecting data sets. Coming up with a concrete plan layout for implementation.

2nd Checkpoint: Nov 10, Complete the first iteration of implementation of complete project and write the midterm report.

3rd Checkpoint: Dec 10, Complete full implementation and complete the mathematical proof of the hypothesis working. Get ready with the presentation and project report.

• **Expectations:** I am currently on the learning curve to Natural Language Processing and would like to master that during the course of the project. I would also like to validate my theoretical understanding practically. On of my key expectation is to learn how to represent a network probabilistically and come up with metrics to validate the working of a model. Some nuances would be to learn working of memory networks and LSTMs.

- **References:**

1. https://research.fb.com/downloads/babi/
2. https://arxiv.org/abs/1503.08895
3. https://arxiv.org/pdf/1502.05698v10.pdf
4. https://arxiv.org/pdf/1410.3916v10.pdf
5. https://arxiv.org/pdf/1506.07285v1.pdf
6. https://papers.nips.cc/paper/5846-end-to-end-memory-networks.pdf
7. https://dl.acm.org/citation.cfm?id=3083294