Chapter 19
Question Answering Chatbot Using Memory Networks

Riddhi Patel
Charotar University of Science and Technology, India

Parth Patel
Charotar University of Science and Technology, India

Chintan M. Bhatt
Charotar University of Science and Technology, India

Mrugendrasinh L. Rahevar
iD https://orcid.org/0000-0002-0551-9229
Charotar University of Science and Technology, India

ABSTRACT

Manipulation of a large amount of data becomes a very tedious task. Hence, the authors took the approach of memory networks for the implementation of the chatbot. Traditionally, the LSTM model was used to implement chatbots and QA systems. But the LSTM failed to store relevant information when given a longer information set. On the contrary, the memory networks have an additional memory component with it. This can help in storing long information for further use which is greatly advantageous for the QA and chatbot systems as compared to LSTM. The authors trained and tested their model over Facebook's bAbi dataset which consists of several tasks and has questions regarding each task to retrieve the accuracy of the model. On the pedestal of that dataset, they have presented the accuracy for every task in their study with memory networks.

INTRODUCTION

The digital age has bombarded mankind with a large number of devices. The devices stand to retrieve information based on the data they are suspected to. Also, give the users the response they need in the large retrieval-based world of a humongous amount of data that resides in the world. The digital devices

rely on information retrieval systems such as a chatbot system. The question answering chatbot system strives to retrieve information from a large repository of data with extreme precision and lesser redundancy. This retrieval system is implemented with the foundation of Artificial Intelligence (AI), (Ainouz, S. A. Ben Ahmed, Mohammed, 2020). Also, a chatbot is more like a human and requires intelligence. Artificial Intelligence techniques are essential in its implementation, (M. Lewkowitz, 2014, Feb 12). One of the techniques to be considered as a part of AI is Machine learning (ML). ML in layman terms can be defined as the ability of a machine to learn on its own from the data it is provided and create a prediction or a decision based on the algorithm that is fed into the machine. The chatbot system also requires techniques to mimic a human brain to generate an accurate response, (Bing Liu1, G. T., Hakkani-Tur, P. S. Heck). That is where Deep Learning comes into the picture showing the neural network similar to nerves in the brain of a human.

A chatbot system can be categorized as a retrieval based system and a rule based system, (Jincy Susan Thomas, Seena Thomas, 2018). The proposed system is a chatbot that uses Memory Networks based approach which makes it better than rule-based systems. This system is a Retrieval based system as the bot is trained on a set of questions and their possible outcomes. Memory networks as the name suggests possess an external memory unlike other Neural networks. Having a memory module becomes extremely necessary as the questions sometimes can tend to be so long that a common neural network would fail to backpropagate entirely. Thus, the external memory module can come into play whenever needed in order to backpropagate and process an entire question, (Sainbayar Sukhbaatar, Arthur Szlam Jason Weston Rob Fergus). The instruction set given to the bot makes it possible to get the answer from the dataset it is trained on inorder to get the most relevant answer and output the same. A good and efficiently pre processed dataset can enable the chatbot to produce new answers. Facebook engineers combined a dataset named bAbi inorder to be used as a task response system. The bAbi set consists of 20 tasks that have variable answers, (Jason Weston, Bordes, S.C., Alexander M. Rush, Bart van Merrienboer, Armand Joulin Tomas Mikolov). We have used this dataset to train our chatbot.

The structure of the chatbot that the authors propose has been illustrated in figure 1. The figure 1 shows that whenever an user asks a question, it does the analysis of the question. Further, it retrieves the necessary document that might have an answer for the question for e.g. "where" questions will have answers in "places" documents. Then it retrieves the answer and analyzes it for it's correctness and finally displays it to the user.

Thus this chatbot has been trained on the machine learning model of memory networks. As, training them end to end requires very less supervision during training which makes it useful in regular scenarios.

BACKGROUND

Evolution of the Chatbot

In 1994, Mauldin coined the term 'ChatterBot' to define the systems aiming to pass the Turing Test. The first ever chatbot to be made in the computer science field is "Eliza". Eliza was developed by Joseph Weizenbaum. Eliza's model captured the inputs and tried to match keywords obtained by rephrasing the input, with predefined responses. The below response can be considered as an example,

utterance = # I want #

Answer = "Why do you want ___?"

Related Content

Artificial Intelligence in Marketing: A Review of Consumer-AI Interactions
Özge Src (2021). *Handbook of Research on Applied Data Science and Artificial Intelligence in Business and Industry (pp. 342-365).*
www.irma-international.org/chapter/artificial-intelligence-in-marketing/284988

An Approach to Supporting Teaching With Data in the Social Sciences
Michele Hayslett, Angela Bardeenand Kayla Olson (2022). *Handbook of Research on Academic Libraries as Partners in Data Science Ecosystems (pp. 1-24).*
www.irma-international.org/chapter/an-approach-to-supporting-teaching-with-data-in-the-social-sciences/302755

Secure Digital Voting System Based on Blockchain Technology
Kashif Mehboob Khan, Junaid Arshadand Muhammad Mubashir Khan (2021). *Research Anthology on Blockchain Technology in Business, Healthcare, Education, and Government (pp. 1280-1290).*
www.irma-international.org/chapter/secure-digital-voting-system-based-on-blockchain-technology/268661

E-Learning as a Training Concept for Staff
Agnieszka Wierzbicka (2024). *Applications of Synthetic High Dimensional Data (pp. 77-93).*
www.irma-international.org/chapter/e-learning-as-a-training-concept-for-staff/342986

Fog Computing and Edge Computing for the Strengthening of Structural Monitoring Systems in Health and Early Warning Score Based on Internet of Things
Leonardo Juan Ramirez Lopezand Gabriel Alberto Puerta Aponte (2022). *Research Anthology on Edge Computing Protocols, Applications, and Integration (pp. 539-563).*
www.irma-international.org/chapter/fog-computing-and-edge-computing-for-the-strengthening-of-structural-monitoring-systems-in-health-and-early-warning-score-based-on-internet-of-things/304323