CREATE CHATBOT USING DEEP LEARNING AND NLP

*Submitted in partial fulfillment of the requirements for the degree of*

*Bachelor of Engineering (SEM VI) Mini Project Synopsis Report - Stage-I*

*by*

SANKET WANKHADE

ROLL NO.59

KARAN JAISWAR

ROLL NO.25

PRITAM TAWADE ROLL NO.53

*Under the Supervision of*



DEPARTMENT OF INFORMATION TECHNOLOGY

KONKAN GYANPEETH COLLEGE OF ENGINEERING

KARJAT-410201

April 2019

# Certificate

This is to certify that the project entitled Automatic Image Annotation is a bonafide work of SANKET WANKHADE (Roll No. 59), KARAN JAISWAR (Roll No. 25), PRITAM TAWADE (Roll No. 53) submitted to the University of Mumbai in partial fulfillment of the requirement for the award of the degree of Undergraduate in DEPARTMENT OF INFORMATION TECHNOLOGY

Supervisor/Guide

Prof.A. Palsodkar

Department of Information Technology

Head of Department Principal

Prof. J.P. Patil Dr. M. J. Lengare

Department of Information Technology Konkan Gyanpeeth College of Engineering

# 

# Project Report Approval

This thesis / dissertation/mini project report entitled Create Chatbot Using Deep

Learning And NLP by Sanket Wankhade (Roll No.59), Karan Jaiswar (Roll No. 25),

Pritam Tawade (Roll No. 53) is approved for the degree of DEPARTMENT OF

INFORMATION TECHNOLOGY

Examiners

1..........................................

2..........................................

Date.

Place.

ii

# Declaration

I declare that this written submission represents my ideas in my own words and where others' ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Signature

Sanket Wankhade (Roll no.59)

Signature

Karan Jaiswar (Roll no.25)

Signature

Pritam Tawade (roll no.53)

Date.

**iii**

# *Abstract*

Today the use of Machine Learning, Artificial Intelligence, Natural Language Processing, and Deep Learning is on the rise. In today’s world using chatbot for customer care services, banking services, credit card companies, and start-ups is common. Using these conversational agents not only eases the work of the companies but also deals with the queries fastly. By using neural network we can train the chatbot to reply as close to as human does. So, by using RNN we can build and train a model accordingly. I have taken the dataset of movie dialogues and used it to build our seq2seq model. By using this dataset and training our model we can build a chatbot that can reply according to the input given by the user. There are many chatbots like code-based or interfaced but they are not good at holding the conversation just like humans. By using the seq2seq model Based on encoder and decoder architecture, we can build a chatbot that can emulate like human beings.

iv

# *Acknowledgements*

Success is nourished under the combination of perfect guidance, care blessing. Acknowledgement is the best way to convey. We express deep sense of gratitude brightness to the outstanding permutations associated with success. Last few years spend in this estimated institution has molded us into condent and aspiring Engineers. We express our sense of gratitude towards our project guide Prof.A.Palsodkar . It is because of his valuable guidance, analytical approach and encouragement that we could learn and work on the project. We will always cherish the great experience to work under their enthusiastic guidance. We are also grateful to our principle Dr . M.J. Lengare who not only supporting us in our project but has also encouraging for every creative activity. We extend our special thanks to all teaching and non-teaching staff, friends and well wishers who directly or indirectly contributing for the success of our maiden mission. Finally, how can we forget our parents whose loving support and faith in us remains our prime source of inspiration . Lastly we would like to thank all those who directly and indirectly helping to complete this project. We would also like to acknowledge with much appreciation the crucial role of the staff of Information Technology Department, who gave the permission to use the all required software/hardware and the necessary material to completing to the project.

V

# Contents

[Certificate i](#_Toc17551)

[Project Report Approval ii](#_Toc17552)

[Declaration iii](#_Toc17553)

[Abstract iv](#_Toc17554)

[Acknowledgements v](#_Toc17555)

[Contents vi](#_Toc17556)

Abbreviations

**MNB** **M**ultinomial **N**aive **B**ayes

**NLP**  **N**atural **L**anguage **P**rocessing

**ML**  **M**achine **L**earning

**AI**  **A**rtifical **I**ntelligence

**RNN** **R**ecurrent **N**eural **N**etwork

**DL** **D**eep **L**earning

Chapter 1

INTRODUCTION

**1.1 Introduction**

Chatbots or Virtual Assistants make the interaction between computers and humans simple. A chatbot uses Artificial Intelligence and can interact with the human in a more realistic way in natural language via various chat rooms, websites, mobile apps, and messaging applications. When any response is generated to user queries in a more realistic human-like language, it is generated via Natural Language Processing. Chatbots provides various services one of which is improving customer experience. In chatbots various crucial role like training, optimizing and configuring the chatbot system is provided by human intervention. The ability of the chatbot to identify the user intent in order to extract relevant entities and generating an adequate response to the User Request is the core function. If any failure occurs to correctly acknowledge the User Request, it will fail to provide the Desired Output. The Response to the User by the chatbot can be predefined generic text, text obtained from a database having different Outputs, Data stored in various Systems.

**1.2 Objectives**

* Chatbots are mainly used to provide customer support.
* Can Schedule meetings, Broadcast newsletters, auto-sequences
* Chatbots are very intelligent. You train them once and they will communicate with your target audience in their language.

## 1.3 Motivation

NLP and deep learning is considered as one of the challenging and growing field. Thus makes it an topic of interest. Since NLP is like teaching a human language to computer and with the help of DL this become more intresting.

## 1.4 Purpose, Scope, and Applicability

### 1.4.1 Purpose

In this fast growing world as technology is changing rapidly management of data become more and more difficult as well as it’s predicted that in next 15 years only 18% of data will be in structured form and remaining in unstrctured from(like message we sent online) therefore there should be technology which can help us in getting data in structured form and this can easily done if we teach computer human language, here comes NLP in play and with the help of deep learning NLP become more versatile.

**1.4.2 Scope**

**NLP in Enterprise AI:**

[*Framework for Applying AI in the Enterprise*](about:blank) explains that this renewed interest in NLP has been triggered by the rise in text data, which in turn has triggered research in advanced AI applications. With this Gartner paper in mind, the enterprise Data Science staff can seize fresh opportunities to combine NLP with [Deep Learning](about:blank) techniques to capture hidden insights from business data very quickly.

**NLP for the Future:**

A significant reward of NLP to businesses is the concept of a smart assistant, which has the potential to transform customer experience, leading to customer loyalty. The smart assistants have already proved their usefulness in customer service, and hopefully NLP will emerge as a [game changer *for CS in the future*.](about:blank)

**The Use in Business Sectors:**

Another remarkable use of NLP may be in sentiment analysis, where texts surrounding social gestures or comments may give a clue to *whether such gestures or comments are positive or negative.* With further improvements in speech recognition technology, the audio-video sources will offer rich data analysis, thus expanding the scope of traditional BI into every aspect of business.

### 1.4.3 Applicability

The system behind the NLP concept is statistical in nature. For this concept to move from Natural Language Processing (NLP) to Natural Language Understanding (NLU) where the consumer can get to see and experience a human emotional connect with the machines, is the future prospect to work towards. Over the last decade, the information technology industry has taken its leap of faith and dug deep into the various aspects of the Natural Language Processing.

Business organisations have found, tested and executed most favorable applications of NLP to advance the progress of Business Intelligence. Yet, the technology needs lots of data and processes in place to understand, analyse and respond to the needs of the human mind.

1. Text Classifiction
2. Language Modeling
3. Caption Generation
4. Speech Recognition
5. Machine Translation
6. Document Summarization
7. Question Answering
8. Sentiment analysis
9. Chatbot
10. Market Intelligence

Chapter 2

LITERATURE SURVEY AND PAPER REVIEW

Literature Survey: Is the process of analyzing, summarizing, organizing, and presenting novel conclusions from the results of technical review of large number of recently published scholarly articles. In this chapter we survey previous research done on NLP and DL, we have studied about following papers published by some experts.

**2.1 Literature Survey**

**1] A Smart Chatbot Architecture based NLP and Machine learning for health care assistance.**

***Author: Soufyane Ayanouz, Mohammed Benhmed, Boudhir Anouar Abdelhakim***

In this authors try to explain importance of chatbot and how we improve it using NLP and ML according to the scientific community, chatbots are user-friendly and any person who has an awareness of typing in their language on the desktop version and in the mobile application can use these chatbots very easily. The new development in artificial intelligence and the new wave of thinking have the potential to entirely change the experience of customers to provide the best services in such a way that echoes with the modern customers. Especially in the field of medicine, a medical based chatbot offers a personalized analysis based upon symptoms. In the future, the recognition of the symptoms of bots and the performance of diagnosis will be highly improved with the addition of support for further medical features, such as symptoms intensity, duration, location, and a more detailed description of symptoms. This study presents state of the art in this field, whichs open us to more exciting works in the future.

## 2] An Intelligent Chat-bot using Natural Language Processing

***Authors: Rishabh Shah, Siddhant Lahoti, Prof. Lavanya. K***

In this authors try to conduct experiment with small dataset with some dynamic variables and different algorithm. It is hard to draw any conclusions at all considering that the amount of user tests conducted was small. With better datasets and knowledge base, better results will be showcased by the bot. But since there are some major components where there have been dynamic changes. The IR (information retrieval) process tends to change with time. Different algorithms have been used and with technology evolving the retrieval process including the triggering part is getting faster. Not just IR but also the answering part with the word embedding technique coming into play makes it easier to relate or connect to different words. Such word embedding makes LSTM network more efficient. The tendencies of the results are promising in that a NLP and self-learning techniques would be a good addition to a chatbot as to make it more humanlike. Although, further research would be required to make any real conclusions.

## 3] BANK CHAT BOT – An Intelligent Assistant System Using NLP and Machine Learning

***Authors: Chaitrali S. Kulkarni, Amruta U. Bhavsar, Savita R. Pingale, Prof. Satish S. Kumbhar.***

In this authors tried 2 different experiment one was to tried different queries to validate the implementation of Bank Chat Bot. In this experiment, they have entered queries which are similar to the questions present in there data set. The analysis of the result is as 87% correct answer and 13% incorrect answer. Second was that they have tried same query in different forms. e.g. Query to open an account in bank can be asked in different ways like:

1. steps for opening account
2. process for opening account
3. how to open an account
4. i want to open account
5. opening account
6. procedure for opening an account in bank

Even if we enter same query in different forms, chat bot can respond same answer accurately to each form.

### 4]Chatbot using NLP and Deep Learning

***Authors: Ravi Khevaria***

Here we see that after training the dataset now the chatbot is tested by running various commands in the terminal. After training the model now the chatbot is ready to be tested. The output generated by the chatbot has moderate relevancy. Many of the outputs can be repetitive and generic. Also due to the lack of real-life quality data, the chatbot performed somehow above average for imitating human interaction. The performance of the chatbot can be optimized by using various real-life datasets. More realistic and high-quality conversational data could further be able to imitate a human being. For future training, personal chat history can also be incorporated to give the chatbot some personality. The replies of the Chatbot can look repetitive and lacks proper relevancy but this can be reduced by adding more diverse and healthy data. The Chatbot developed using the seq2seq model can be further improved with more robust and high-quality real-life datasets which can further increase the nature of the replies of the chatbot. To further advance the project, Deep Reinforcement Learning can be applied that could improve the results significantly. The main techniques include Deep Neural Network, Recurrent Neural Network, seq2seq modeling with encoder and decoder. Future techniques to improve can include Long Short Term Memory Based RNN cells, Bi-directional LSTM, Neural Attention Model and Beam Searching. The training on Cornell Corpus produced results that can further be improved by more attention and speculation on training parameters. Training with other hyper-parameters and different datasets for further experimentation.

### 5] Implementation of a Chatbot System using AI and NLP

***Authors: Tarun Lalwani, Shashank Bhalotia, Ashish Pal, Shreya***

***Bisen,Vasundhara Rathod***

In this authors try to creates a chatbot for the college purpose to solve the difficulty such as, on a college’s website, one often doesn’t know where to search for some kind of information. It becomes difficult to extract information for a person who is not a student or employee there. The solution to these comes up with a college inquiry chat bot, a fast, standard and informative widget to enhance college website’s user experience and provide effective information to the user. Chat bots are an intelligent system being developed using artificial intelligence (AI) and natural language processing (NLP) algorithms. It has an effective user interface and answers the queries related to examination cell, admission, academics, users’ attendance and grade point average, placement cell and other miscellaneous activities. It is often impossible to get all the data on a single interface without the complications of going through multiple forms and windows. The college chatbot aims to remove this difficulty by providing a common and user-friendly interface to solve queries of college students and teachers. The purpose of a chatbot system is to simulate a human conversation. Its architecture integrates a language model and computational algorithm to emulate information online communication between a human and a computer using natural language.

The college student and employees can freely upload their queries. The chatbot provides fast and efficient search for answers to the queries and gets the relevant links to their question. A background research took place, which included an overview of the conversation procedure and tries to find out the relevant keywords related to that query to provide the proper link. The database storage includes information about questions, answers, keywords, and logs. We have also developed an interface. The interface developed will have two parts, one for users and the other for the administrator

**2.2 Paper Comparison**

|  |  |  |  |
| --- | --- | --- | --- |
| **Sr.no** | **Paper Title** | **Authors** | **Description** |
| **1** | **A Smart Chatbot Architecture based NLP and Machine learning for health care assistance.** | **Soufyane**  **Ayanouz,**  **Mohammed**  **Benhmed,**  **Boudhir**  **Anouar**  **Abdelhakim** | explain importance of chatbot and how we improve it using  NLP and ML according to the scientific community, chatbots are user-friendly and any person who has an awareness of typing in their language on the desktop version and in the mobile application can use these chatbots very easily |
| **2** | **An Intelligent Chat-bot using Natural Language Processing** | **Rishabh Shah,**  **Siddhant**  **Lahoti,**  **Prof. Lavanya. K** | experiment with small dataset with some dynamic variables and different algorithm, Different algorithms have been used and with technology evolving the retrieval process including the triggering part is getting faster. |
| **3** | **BANK CHAT BOT – An Intelligent**  **Assistant System Using NLP and**  **Machine Learning** | **Chaitrali S. Kulkarni,**  **Amruta U.**  **Bhavsar, Savita**  **R. Pingale, Prof. Satish S.**  **Kumbhar** | They have entered queries which are similar to the questions asked while creating bank accounts. The analysis of the result is as 87% correct answer and 13% incorrect answer |
| **4** | **Chatbot using NLP and Deep Learning** | **Ravi Khevaria** | Training the dataset now the chatbot is tested by running various commands in the terminal. After training the model now the chatbot is ready to be tested. The output generated by the chatbot has moderate relevancy |
| **5** | **Implementation of a Chatbot System using AI and NLP** | **Tarun Lalwani,**  **Shashank**  **Bhalotia,**  **Ashish Pal,**  **Shreya Bisen,**  **Vasundhara**  **Rathod** | The purpose of a chatbot system is to simulate a human conversation. Its architecture integrates a language model and computational algorithm to emulate information online communication between a human and a computer using natural language |

Figure 2.1Comparison table

Chapter 3

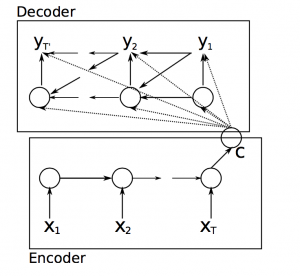
SURVEY OF METHODOLOGY

**3.0.1 Deep Learning**

The field of machine learning is witnessing its golden era as deep learning slowly becomes the leader in this domain. Deep learning uses multiple layers to represent the abstractions of data to build computational models. Some key enabler deep learning algorithms such as generative adversarial networks, convolutional neural networks, and model transfers have completely changed our perception of information processing. However, there exists an aperture of understanding behind this tremendously fast-paced domain, because it was never previously represented from a multi scope perspective. The lack of core understanding renders these powerful methods as black-box machines that inhibit development at a fundamental level. Moreover, deep learning has repeatedly been perceived as a silver bullet to all stumbling blocks in machine learning, which is far from the truth. This article presents a comprehensive review of historical and recent state-of-the-art approaches in visual, audio, and text processing; social network analysis; and natural language processing, followed by the in-depth analysis on pivoting and ground breaking advances in deep learning applications. It was also undertaken to review the issues faced in deep learning such as unsupervised learning, black-box models, and online learning and to illustrate how these challenges can be transformed into prolific future research avenues.

**3.0.1. a Sequence to Sequence modelling technique**

A typical sequence to sequence model has two parts – an **encoder** and a **decoder.** Both the parts are practically two different neural network models combined into one giant network.



Broadly, the task of an encoder network is to understand the input sequence, and create a smaller dimensional representation of it. This representation is then forwarded to a decoder network which generates a sequence of its own that represents the output.

3.0.2 Natural Language Processing

NLP is the field of research and application that explores how computers can be used to understand and manipulate natural language text or speech to do useful things. NLP can be defined as the automatic processing of human language. Alternate terms that are often preferred are, ‘Language Technology’ or ‘Language Engineering’. It is closely related to linguistics. To summarize, NLP is a discipline which is concerned with the interaction between natural human languages and computing devices. It is a way in which computers analyze, understand and derive meaning from human language in a smart and useful way. This human-computer interaction enables real-world applications like Search engines (Google), Translation systems (Google Translate), automated question answering, Text categorization, Spelling and Grammar checking and more. NLP is integrated widely in a large number of contexts such as evaluation systems, e-learning, research, machine translation, multilingual and cross-language information retrieval (CLIR), speech recognition. As computers play a major role in preparation, storage, analysis and transfer of information, endowing them with the ability to understand and to generate information expressed in natural language becomes important. Hence, the goal of NLP is to design and build computer systems which have the ability to analyze natural languages (like English, German, etc.) and to generate output in natural language. Human linguistic communication occurs as speech and as written language.

**Chapter 4**

REQUIRMENTS AND ANALYSIS

**4.1 Problem Definition**

We aim to create a chat bot using Deep learning and NLP.It is a question answering system which will try to answer the questions which will put across in the form of query. So,the question can be definition questions, biographical question & multilingual question among other types of question asked in natural languages are answer by such systems . Q& Ans. A is well researched problem in NLP. So using DL we are going to solve basic NLP tasks.

**4.2 Requirements Specification**

For implementation, in software we will require the following software and hardware specifications:

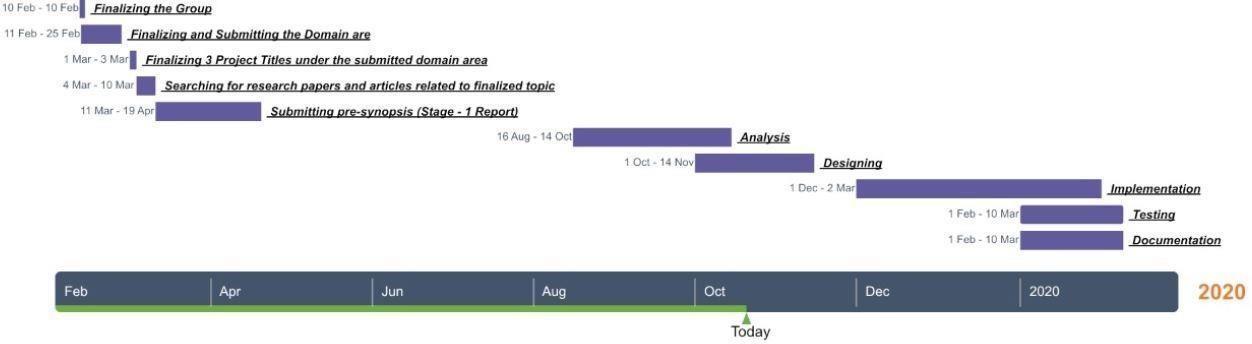
## • Software Specification

For implementation we will require applications such as MAT lab, Jupiter Notebook, Python IDE, etc. That can be used to build an environment based on neural network and to train machine based on it.

## • Hardware Specification

To implement the project we will require a computer with specification such as multicore CPU, graphics card, hard disk upto 500GB, upto 8GB RAM. Input devices such as keyboard, optical mouse. Output device: LCD monitor**.**

## Planning and Scheduling



**Chapter 5**

SYSTEM DESIGN

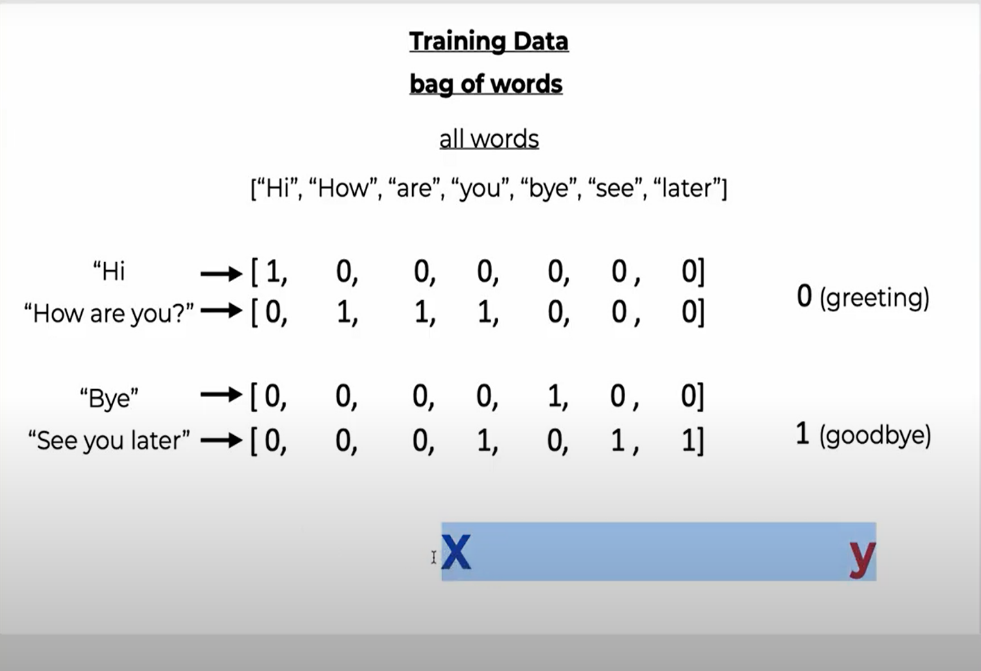


Figure 5.1 Training Data

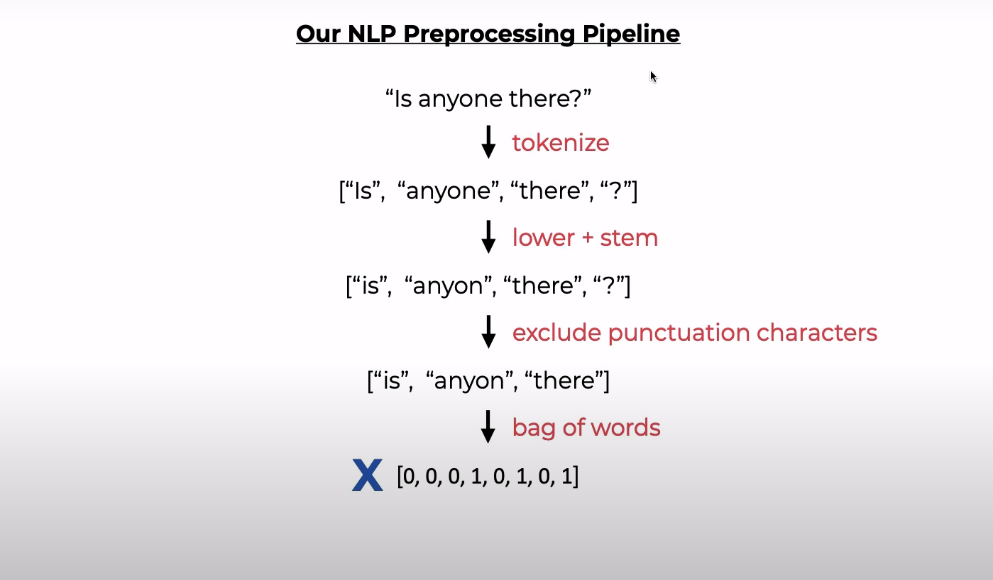


Figure 5.2 NLP Preprocessing Pipeline

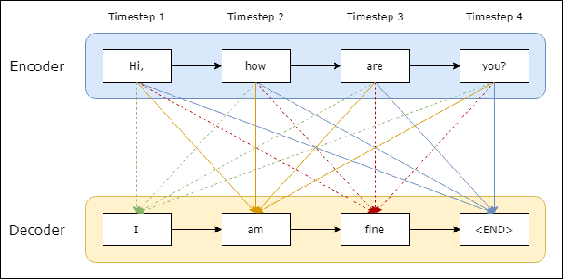
****

Figure 5.3 NLP sequence to sequence model (Refer from google)

**Chapter 6**

CONCLUSIONS

**6.1 Conclusion**

We developed a chatbot using DL and NLP that will answer to user submitted questions. We have used some basic concepts of both NLP (Natural Language Processing) and DL (Deep Learning). First we used a raw data, that is a group of sentences and after cleaning the data we applied our DL models such as Sq2Sq which provides the expected results. So in this way we accomplished the chatbot using Deep Learning and NLP.

BIBILOGRAPHY

[https://chatbotsmagazine.com/contextual-chat-bots-with-tensorflow-4391749d0077](about:blank)

[https://sloboda-studio.com/blog/how-to-use-nlp-for-building-a-chatbot/](about:blank)