ARTIFICIAL INTELLIGENCE WITH REAL LIFE INTERACTION (AIRI)

BY

Amit Kumar Singha

ID: 161-15-7366

MD. Sadakul Islam

ID: 161-15-7377

AND

Redoy Chandraw Das

ID: 161-15-7368

This Report Presented in Partial Fulfillment of the Requirements for the Degree of Bachelor of Science in Computer Science and Engineering

Supervised By

Shah Md. Tanvir Siddiquee
Assistant Professor
Department of CSE
Daffodil International University

Co-Supervised By

Ahmed Al Marouf
Lecturer
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

DECEMBER 2019

APPROVAL

This Project/internship titled **"Artificial Intelligence Based Real-life Interaction"**, submitted by Amit Kumar Singha, ID No: 161-15-7366, MD. Sadakul Islam, ID No: 161-15-7377 and Redoy Chandraw Das, ID No: 161-15-7368 to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on 07th December,2019.
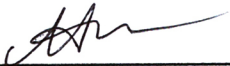
BOARD OF EXAMINERS

Dr. Syed Akhter Hossain   **Chairman**
Professor and Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Nazmun Nessa Moon   **Internal Examiner**
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Gazi Zahirul Islam   **Internal Examiner**
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Dr. Mohammad Shorif Uddin   **External Examiner**
Professor
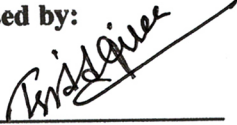Department of Computer Science and Engineering
Jahangirnagar University

DECLARATION

We hereby declare that, this project has been done by us under the supervision of **Shah Md Tanvir Siddiquee, Assistant Professor, Department of CSE** Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.
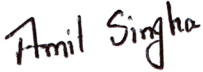
Supervised by:

Shah Md Tanvir Siddiquee
Assistant Professor
Department of CSE
Daffodil International University

Co-Supervised by:

Ahmed Al Marouf
Lecturer
Department of CSE
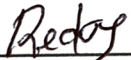Daffodil International University

Submitted by:

Amit Kumar Singha
ID: -161-15-7366
Department of CSE
Daffodil International University

Md. Sadakul Islam
ID: -161-15-7377
Department of CSE
Daffodil International University

Redoy Chandraw Das
ID: -161-15-7368
Department of CSE

ACKNOWLEDGEMENT

First we express our heartiest thanks and gratefulness to almighty God for His divine blessing makes us possible to complete the final year project/internship successfully.

We really grateful and wish our profound our indebtedness to **Mr. Shah Md Tanvir Siddiquee, Assistant Professor**, Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of "*Artificial intelligence*" to carry out this project. His endless patience ,scholarly guidance ,continual encouragement , constant and energetic supervision, constructive criticism , valuable advice ,reading many inferior draft and correcting them at all stage have made it possible to complete this project.

We would like to express our heartiest gratitude to Prof. Dr. Syed Akhter Hossain, Head**,** Department of CSE, for his kind help to finish our project and also to other faculty member and the staff of CSE department of Daffodil International University.

We would like to thank our entire course mate in Daffodil International University, who took part in this discuss while completing the course work.

Finally, we must acknowledge with due respect the constant support and patients of our parents.

ABSTRACT

Around 88% people are using android system in present world. Life without a smart phone is quite impossible in this generation. So, we feel the need of a system that can perform according to users need. Where users don't need to give command by himself. It will help a human being to save his time and perform all smart phone activities while doing other things. Also we build this system that take Bangla command, till currently nobody has enforced a voice assistant which will acknowledge Bengali speech. As a result, our voice assistant will perceive commands given in each English and Bengali language and may perform action in keeping with the command. This distinctive feature can dissent our application from those existing applications as a result of until currently no corporations has enforced Bengali language in their voice assistant.

TABLE OF CONTENTS

CONTENTS	PAGE

LIST OF FIGURES

CHAPTER 1
INTRODUCTION

1.1 Introduction

In today's world, science & technology is growing very fast, and we are getting introduce with different new technologies day by day. And, one of the advance technologies of computer science is Artificial Intelligence (AI) that's is ready to create a new revolution in the world by making intelligent machines and computer programs. Artificial intelligence is a reflection of human insight forms by machines, different PC frameworks.

This technology is a fascinating concept especially now that technic are creating systems that can think and do things almost like humans. This procedure incorporate getting the learning, thinking, and personal growth. Artificial Intelligence is a study of how human brain think, learn, decide then work, when it try to solve problems. And finally this study outputs intelligent software systems. So we also try to build an application that Interact with our daily life. Like News suggestion with user choice, maintain daily ToDo task and update report, Real time voice commend to get any feature of our application. Along the way, we also hope to excite you about the numerous applications and huge possibilities in the field of AI, which continues to expand human capability beyond [1] our imagination.

1.2 Motivation

"If you talk to a man in a language he understands, that goes to his head. If you talk to him in his own language, that goes to his heart." — Nelson Mandela [2]

Around 88% people are using android system in present world. Life without a smart phone is quite impossible in this generation. From SIRI to self-driving cars, AI (AI) is progressing speedily. Till currently no one has actual a voice partner which will understand Bengali discourse.

Thus, our voice associate will comprehend directions given in each English and Bengali language and may perform the activity as indicated by the order. This special part can

distinction our application from those current applications on the grounds that until currently no organizations have actual Bengali language in their voice hand. So, we feel the need of a system that can perform according to users need. Where users don't need to give command by himself. It will help a human being to save his time and perform all smart phone activities while doing other things. Also it will take command by voice command and execute operation upon user choice. So we think we also need to build an AI that perform all user recommended task and take operation giving few specific command. So we need to build AIRI.

1.3 Rationale of the Study

There are many Artificial Intelligence system present in this 21st century. From SIRI to self-driving cars, AI (AI) is progressing speedily. Every AI has his own few feature, they all take command in English language and do specific task on specific platform. But none of AI take any Bangla command, that's why we think that we need to build a AI that take Bangla command and do others task as well. That's why we build AIRI for especially Bangladeshi people where user can command by Bangla voice and also give command manually.

1.4 Research Questions

- ➢ What is an Artificial Intelligence and How it being helpful?
- ➢ How the existing system and what is the problem?
- ➢ What is the background knowledge necessary to fully understand this problem?
- ➢ How it will solve the existing problem?
- ➢ How you add Bangla voice command?
- ➢ Is this system takes command from user on every step?
- ➢ What were the results of these solutions?

- ➢ Is there a cost involved (salaries/space/equipment/supplies/etc)? If so, how would it be funded?
- ➢ Why is the problem so challenging to solve?

- ➢ Prove that a particular method is more efficient than other methods?
- ➢ How you detect voice command?

1.5 Expected Output

We are hoping to see the following outputs:

- ➢ Bangla Voice Commend
- ➢ Bangla recognize
- ➢ Text Analysis
- ➢ Realize users are awake or sleeping using sensor.
- ➢ Alarm, Weather & To-do task.
- ➢ Give news according to the choices.
- ➢ Show nearest suggestion of any (like hospital, bank, and restaurant).
- ➢ Important contact number.

CHAPTER 2

BACKGROUND

2.1 Introduction

We are living in an era where technology advancement is increasing day by day [3]. In this present time a lot of work is taken over by machines, software and various automatic processes [4]. In all the advancement artificial intelligence is the most popular creation. AI is the science which is used to develop a machine like human brains. It can be used to do some works that human needs to do on regular basis. It can actually make our daily life advanced by using modern technology. It is efficient enough to reduce human efforts in various areas [5]. AI is used in different sectors such as banking sectors, medical science, air transport, gaming, heavy industry, entertainment and more. Now a day there are so many companies have already launched their voice assistant feature as a part of artificial intelligence. They are Google Assistant, Apple Siri, Amazon Alexa, Samsung Bixby, Microsoft Cortana etc.

The use of artificial intelligence is increasing with the passage of time. So we think to make something that can help user to use AI easily in an android application. Many giant companies has already created this type of features. We have tried a little different thing from them. The unique part of our application is that we have a voice assistant in our application that can hear command from the user in both English and Bengali. By hearing the command the voice assistant can execute the command. Language selection is user choice which one he or she prefers. We have named our voice assistant as AIRI.

In our application we have some common features that already exists in other applications. We have tried to combine all those applications in one place and connect them with the voice assistant so that user can use all the features through voice command. This will reduce user effort and user can use the application easily. Our intention is to make our application as user friendly.

2.2 Related Work

The application AIRI that we have implemented is related with other applications those are implemented by giant companies. We are influenced by those existing applications. We have tried to implement something unique that differs from the existing applications. Artificial Intelligence is the common factor in all the application. In present time, voice assistant is the most popular creation of artificial intelligence. The voice assistant that already exists it can understand commands given in English and other foreign languages but it does not recognize Bengali language. So we tried to implement a voice assistant that is able to recognize Bengali language. Till now no one has implemented a voice assistant that can recognize Bengali speech. As a result, our voice assistant can understand commands given in both English and Bengali language and can perform action according to the command. This unique feature will differ our application from those existing applications because till now no companies has implemented Bengali language in their voice assistant.

2.2.1 Google Assistant

Google assistant is an artificial intelligence developed by google. It is available in all smart phone devices that supports Android and IOS. It is also available in smart home devices. User can interact with google assistant through voice command or keyboard input. Google assistant is able to search the internet, schedule event and alarm, adjust hardware settings on user's device, make memory of photos, gives real time answer, control home and show information from the user's Google account. Google has also announced that the assistant will be able to identify objects, gather visual information through the device's camera, identify songs and as well [6] as purchasing products and send money. Google assistant can be enabled by the user by saying, "Ok Google". The google assistant allows a user to activate and modify vocal shortcut commands in order to perform actions on their device. The feature of speech recognition is available in English and other languages. It is one of the most advanced and dynamic assistant. The user base of Google assistant is huge.

2.2.2 Apple Siri

Siri is a virtual assistant that is a part of Apple IOS and it is developed by Apple. The assistant uses voice queries and natural language user interface to answer questions. Siri is a spin off from a project originally developed by SRI International Artificial Intelligence Center. Its speech recognition engine was provided by Nuance Communication and Siri uses advanced machine learning technology. Initially it was released as an app for IOS [7]. Then the separate app was removed and Siri was included as a part of Apple's product. The assistant is able to make calls or send texts and can set alarm, timer, reminder, get direction, preview calendar. It can also find songs according to user preference. It can translate words and phrases, handle payment, handle device settings, and search on the internet. It has a feature called Shortcuts, it comes up on lock screen or search result. Siri can be enabled by the user by saying, "Go Siri" and answers all kinds of questions. Siri does its job faster. Siri always learning how to be more helpful. Siri keeps user's information private and secure.

2.2.3 Amazon Alexa

Amazon Alexa is simply known as Alexa, is a virtual assistant which is developed by Amazon [8]. This assistant was first used in Amazon Echo and Amazon Echo Dot. It is capable of voice interaction, music playback, making to-do lists, setting alarm, streaming podcasts, playing audiobooks, make phone calls and send message etc [9]. It can also provide news, weather, traffic, sports and other real time information. Alexa can control several smart devices and can be used itself as a home automation system. Users can extend the capabilities of Alexa by installing "skills" [10]. To start up the assistant there is a wake-word called "Alexa". So users do not need to push any button start up the assistant. Users can interact and communicate with Alexa in English and other foreign languages. The new Echo Studio became the first smart speaker with 360 sound and Dolby sound. Amazon has announced a new integration with answer engine that will be able to answer question related to math, science, astronomy, engineering, geography, history and more.

Alexa supports a multitude of subscription-based and free streaming services on Amazon devices [11]. It is able to stream media and music directly through Amazon music library. Alexa is making itself more secure and private. It can erase specific question or command as user choice.

2.2.4 Samsung Bixby

Bixby is a virtual assistant developed by Samsung Electronics. This voice assistant was first introduced in Galaxy S8 series. Like other giant companies Samsung decided to launch their own voice assistant named as "Bixby". It comes with three parts, known as "Bixby Voice", "Bixby Vision", "Bixby Home". It can be activated by calling "Bixby" or press the Bixby button. Bixby button is reprogrammable and could be set to open other application or assistant such as Google Assistant. It is designed to make available on any other device. It understands natural language and can interpret incomplete information and take action [12]. User can get it to action without laboriously detailing exactly, it can understand what you are trying to say and then it takes the logical next step. The wake word for this assistant is "Hi Bixby". It understands English and few other languages. Bixby voice assistant is able to send message, take selfie, take screenshot, navigate location, weather update, remind schedule, check sport scores, turn down screen brightness, check your calendar, launch apps, read out latest messages, can speak in male or female voice, create album of photos and more. Bixby voice also supports quick command. It allows user to perform multiple actions with a single phrase. Such as if user wants to open slow-motion camera, user can just command as slow-motion and the assistant opens it directly. Bixby understands thousands of commands. It makes the user information secure and private. Samsung said that Bixby is faster at responding queries.

2.2.5 Microsoft Cortana

Cortana is a virtual assistant created by Microsoft [13]. It has been launched as a key ingredient of Microsoft's planned "makeover" of the future operating systems for

Windows phone and Windows [14]. It was named Cortana as a synthetic intelligence character in Microsoft from the popular Halo game franchise.

Cortana is Microsoft's first major foray into consumer facing artificial intelligence. Microsoft has intentionally built Cortana having a long term vision [15]. The goal is to support all types of human interaction. The wake word for Cortana is "Hey, Cortana" [16]. It is able to set reminder, recognize natural voice without using keyboard input, answer question, compose an email, multi device syncing, multitasking, offers personalize recommendation and more. It is powered by "Bing". Through this Cortana gets to know the user and builds a relationship with the user. It gets better over time by asking questions based on user behavior. User will be able to launch Cortana from the start screen. Without Microsoft's Bing platform Cortana would not exist. Cortana is currently available in English and few more languages. It stores personal information in a "Notebook". It is designed in such a way that it seems to indicate activities such as searching or talking. It has a "do not disturb" mode in which user can specify quiet hours. The information stored in Cortana can be deleted by the user. In different regions Cortana uses different voice. It can search files from the storage for user whatever the user commands.

2.3 Research Summary

In this research, our motive is to implement a voice assistant that understands Bengali language. In Bangladesh, most of the people prefers Bengali language than English. Thinking of the comfort zone of our people we have implemented a voice assistant that understands Bengali. This will help our people to use this application easily.

More and more companies are seeking their attention in artificial intelligence and many giant companies has already launched their voice assistant. The assistants are able to understand English and other foreign language. But no one is concern about our mother tongue Bengali language. For this reason, we have made our decision to implement an application using artificial intelligence that contains a voice assistant and this assistant will be able to understand both Bengali and English language.

In our application, user can have other application's features at one place so user do not need to use any other application containing the features we have implemented.

The voice assistant called "AIRI", is able to show daily news, weather, and search from internet, to-do task and more. This application is easy to use. There is a login system to enter into the application. User needs to sign up first then login to the application. Then the user can use all the features available in the application. We concern about user's data and information security. Finally, our application is user friendly and users will feel comfortable and secure using this application

2.4 Scope of the Problem

A voice assistant is implemented to do such things that a user has to perform on a regular basis. Though voice assistant is very helpful it has some scope of problems. Voice assistants has two specific branches voice recognition and natural language processing [24]. When a user gives a command, the voice recognition part converts the sound wave into written words. Then the other branch takes the word and processes the command.

2.4.1 Ambiguity

Voice assistants needs to get better with the ambiguity. The lacking of ambiguity can suffer a user in using voice assistant. For example, a user gives command to the assistant to "wake me up at 10". In this case user do not need to say a.m. or p.m. because voice assistant should determine this based on the time of day. Things like this needs to be determined by the assistant that a user feel comfort using the application.

2.4.2 Privacy and Ethical Issues

Most of the voice assistants do not know who is talking to them. They need to be trained that they can recognize the voice of owner. Currently Google Assistant and Amazon Alexa have the feature of recognizing voice. This is a serious issue because anyone can give command to the voice assistant if it is listening. Voice assistants should resolve this problem otherwise, it will hamper the original user.

2.4.3 Social Impact

As technologies are growing rapidly it affects our behavior in some way personally and socially. Because of the technology people are getting addicted with various devices. They do not have time to interact with someone personally but they interact virtually. The development of voice technology might lead to a point where people will consider reading and writing are the things from the past. Eventually people will stop taking with one another directly and will listen to the recordings instead.

2.5 Challenges

To develop a voice assistant developer faces a lot of challenges. They need to overcome with those challenges. Voice assistants should be trained how it should perform action when a user gives command. It needs to understand user's requirements and language. Most of the existing voice assistants can understand English and other foreign languages. But they are not trained in Bengali language. If a user from Bangladesh who do not know English, wants to give command to the voice assistant then the assistant would not be able to understand the command of the user and it will not be able to perform any action. So developer should train voice assistants in all existing languages all over the world. The first challenge that comes in mind is that ambiguity of a voice assistant. A voice assistant should be smart enough to understand what the user wants to say and perform the action according to the command. The second one is privacy and ethical issues.

Voice assistants must recognize the voice of the actual user and perform action after recognizing the voice command. The third one is social impact. The rapid grow of technology hampers our society in many ways. So we need to be very careful that the technology does not hamper our social environment. Finally, a voice assistant is very helpful for us it shows us the future of modern world. Though there are so many challenges implementing a voice assistant it needs to be perfect for necessary use. By fixing all the challenges we can have the best voice assistant.

CHAPTER 3
RESEARCH METHODOLOGY

3.1 Introduction

To get the voice command and making the decision from the percept of AI, we have to get the clear voice command from the user. Fig 3.1 refers how our system works properly. We use RNN algorithm for processing English to Bangla conversion. Then, feature extractions are prepared from the segmented model. RNN is employed to train the dataset and seq2seq classify the language as output.

A machine learning approach that is described in this research is needed to achieve such a machine vision system, and figure 3.1 shows the steps to build the framework for the process.
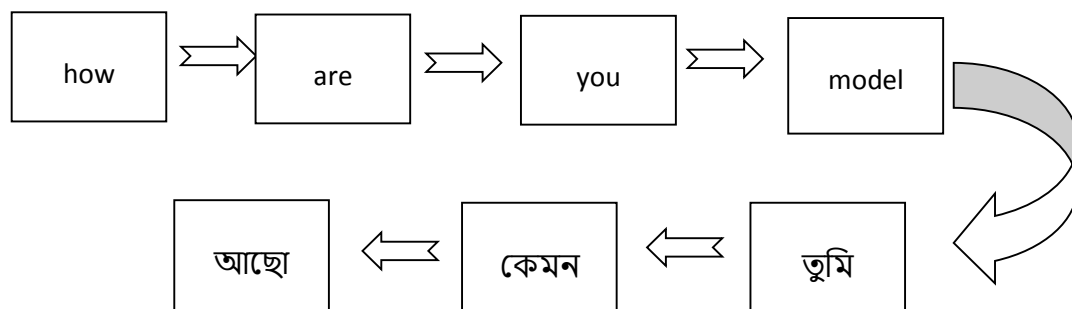
Fig. 3.1.1 Proposed conceptual diagram

Encoder

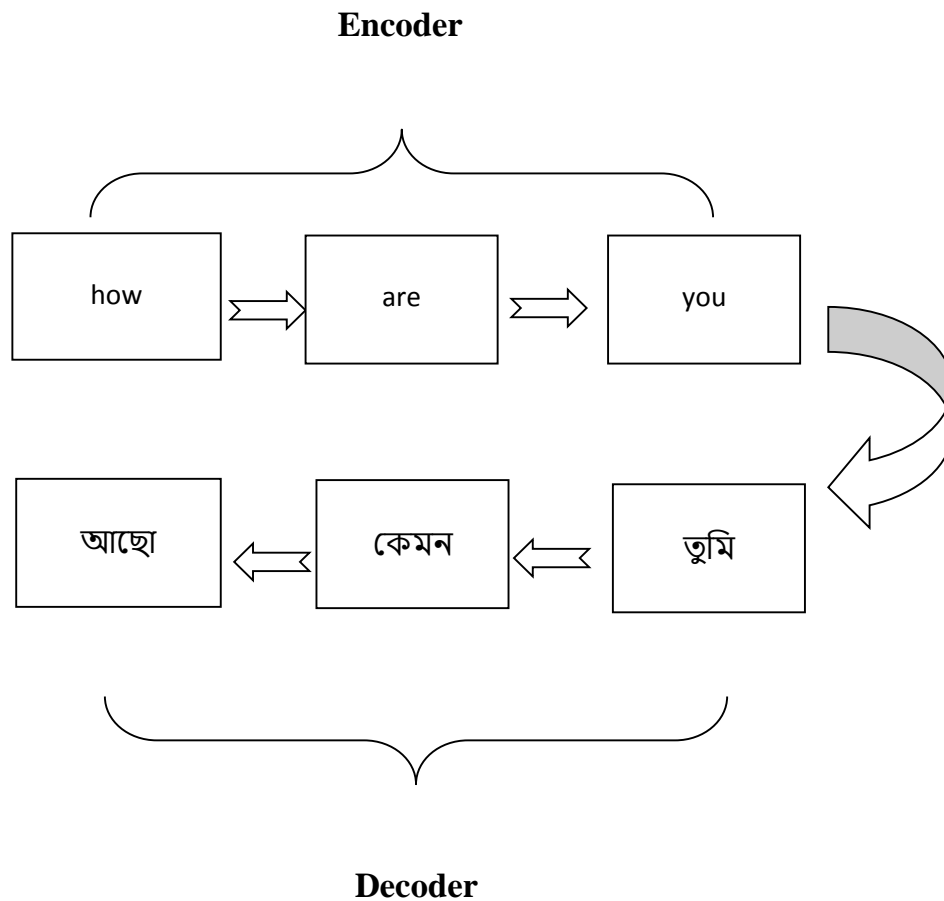Fig. 3.1.2 Proposed conceptual diagram

3.2 Research Subject and Instrumentation

3.2.1 Understanding the Problem Statement

Let's circle back to the place we left off in the introduction section, i.e., gaining knowledge of Bangla. However, this time around we are going to make my machine do this task. The goal is to convert a Bangla sentence to its English using a Neural Machine Translation (NMT) system.

3.2.2 Introduction to Sequence-to-Sequence (Seq2Seq) Modeling

Sequence-to-Sequence (seq2seq) models are used for a variety of NLP tasks, such as text summarization, speech [17] recognition, DNA sequence modeling, among others. Our purpose is to translate given sentences from one language to a different. Here, both, the input and output square measure sentences. In different words, these sentences square measure a sequence of words getting in and out of a model.

This is the core idea of Sequence-to-Sequence modeling [18]. This approach is illustrated in the figure below. A typical seq2seq model has 2 major components

- ➢ an encoder
- ➢ a decoder

Both these components are genuinely two specific recurrent neural network (RNN) models mixed into one massive network:

We've listed a few significant use cases of Sequence-to-Sequence modeling below (apart from Machine Translation, of course):

- ➢ Recognition of speech
- ➢ Name Entity / Subject Extraction for distinguishing the main subject from a text body
- ➢ Relation Classification to tag relationships between various entities tagged in the above step
- ➢ Chatbot skills to have conversational ability and engage with customers
- ➢ Text Summarization to generate a concise summary of a large amount of text

3.2.3 Implementation in Python using Keras

It is time to urge our hands dirty! there's no higher feeling than learning a subject by seeing the results first-hand. We will fire up our favorite Python environment (Jupyter Notebook for me) and get straight down to business.

3.2.4 Import the Required Libraries

```
1   import string
2   import re
3   from numpy import array, argmax, random, take
4   import pandas as pd
5   from keras.models import Sequential
6   from keras.layers import Dense, LSTM, Embedding, RepeatVector
7   from keras.preprocessing.text import Tokenizer
8   from keras.callbacks import ModelCheckpoint
9   from keras.preprocessing.sequence import pad_sequences
10  from keras.models import load_model
11  from keras import optimizers
12  import matplotlib.pyplot as plt
13  % matplotlib inline
14  pd.set_option('display.max_colwidth', 200)
```

Fig. 3.2.4.1. Import Libraries

3.2.5 Read the Data into our IDE

Our data is a text content file (.txt) of English-Bangla sentence pairs. First, we use the function described below to read the file.

```
1   # function to read raw text file
2   def read_text(filename):
3           # open the file
4           file = open(filename, mode='rt', encoding='utf-8')
5
6           # read all text
7           text = file.read()
8           file.close()
9           return text
```

Fig. 3.2.5.1 Read data from text file

Let's define another function to split the text into English-Bangla pairs separated by '\n'. We'll then split these pairs into English sentences and Bangla sentences severally.

```
1   # split text into sentences
2   def to_lines(text):
3           sents = text.strip().split('\n')
4           sents = [i.split('\t') for i in sents]
5           return sents
```

Fig.3.2.5.2 Split the text

We can now use these features to examine the text into an array in our desired format.

```
data = read_text("deu.txt")
deu_eng = to_lines(data)
deu_eng = array(deu_eng)
```

Fig. 3.2.5.3 Text into array

The actual data contains over 150,000 sentence-pairs. However, we will use only the first 50,000 sentence pairs to reduce the training time of the model.
You can change this number as per your system's computation power (or if you're feeling lucky!).

```
deu_eng = deu_eng[:50000,:]
```

Fig. 3.2.5.4 Used pairs

3.2.6 Text Pre-Processing

Quite an essential step in any project, especially so in NLP. The data we work with is extra regularly than not unstructured so there are certain things we want to take care of before jumping to the mannequin building part.

(a) Text Cleaning

Let's first take a look at our data. This will help us determine which pre-processing steps to adopt.

```
deu_eng
```

```
array([['hi', 'hallo'],
       ['hi', 'grüß gott'],
       ['run', 'lauf'],
       ...,
       ['mary has very long hair', 'maria hat sehr langes haar'],
       ['mary is toms secretary', 'maria ist toms sekretärin'],
       ['mary is a married woman', 'maria ist eine verheiratete frau']],
      dtype='<U380')
```

Fig. 3.2.6.1 Show pair

We will get rid of the punctuation marks and then convert all the text to lower case.

```
1   # Remove punctuation
2   deu_eng[:,0] = [s.translate(str.maketrans('', '', string.punctuation)) for s in deu_eng[:,0]]
3   deu_eng[:,1] = [s.translate(str.maketrans('', '', string.punctuation)) for s in deu_eng[:,1]]
4
5   # convert text to lowercase
6   for i in range(len(deu_eng)):
7       deu_eng[i,0] = deu_eng[i,0].lower()
8       deu_eng[i,1] = deu_eng[i,1].lower()
```

Fig. 3.2.6.2 Remove punctuation

(b) Text to Sequence Conversion

A Seq2Seq model requires both input and output sentences to be converted into fixed-length integer sequences.

But let's visualize the length of the sentences before we do that [19]. The lengths of all sentences will be recorded in two separate lists of English and Bangla.

```
1    # empty lists
2    eng_l = []
3    deu_l = []
4
5    # populate the lists with sentence lengths
6    for i in deu_eng[:,0]:
7           eng_l.append(len(i.split()))
8
9    for i in deu_eng[:,1]:
10          deu_l.append(len(i.split()))
11
12   length_df = pd.DataFrame({'eng':eng_l, 'deu':deu_l}) length_df.hist(bins = 30)
13   plt.show()
```

Fig. 3.2.6.3 Convert into integer Sequence

Fig. 3.2.6.4 Showing integer sequence
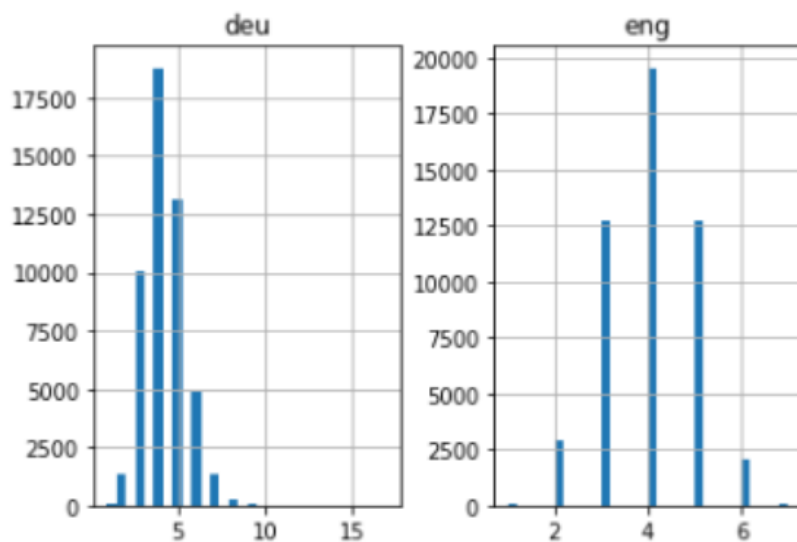
Quite intuitive — the maximum length of the Bangla sentences is 11 and that of the English phrases is 8.

Next, vectorize our text data by using Keras's Tokenizer() class. It's going to turn our sentences into integer sequences. Then we can pad those sequences with zeros to make all the same length sequences [20].

Note that both Bangla and English sentences will be prepared for tokenizers:

```
1    # function to build a tokenizer
2    def tokenization(lines):
3            tokenizer = Tokenizer()
4            tokenizer.fit_on_texts(lines)
5            return tokenizer
6
7    # prepare english tokenizer
8    eng_tokenizer = tokenization(deu_eng[:, 0])
9    eng_vocab_size = len(eng_tokenizer.word_index) + 1
10   eng_length = 8
11
12   # print('English Vocabulary Size: %d' % eng_vocab_size)
13
14   # prepare Deutch tokenizer
15   deu_tokenizer = tokenization(deu_eng[:, 1])
16   deu_vocab_size = len(deu_tokenizer.word_index) + 1
17   deu_length = 8
18
19   # print('Deutch Vocabulary Size: %d' % deu_vocab_size)
```

Fig. 3.2.6.5 Token generation

The below code block incorporates a feature to prepare the sequences. It will additionally perform sequence padding to a maximum sentence size as stated above.

```
1    # encode and pad sequences
2    def encode_sequences(tokenizer, length, lines):
3            # integer encode sequences
4            seq = tokenizer.texts_to_sequences(lines)
5            # pad sequences with 0 values
6            seq = pad_sequences(seq, maxlen=length, padding='post')
7            return seq
```

Fig. 3.2.6.8 Encoding Sequence

3.2.7 Model Building

We are now dividing the data into train and test set, respectively, for model training and evaluation.

```
1    from sklearn.model_selection import train_test_split
2
3    # split data into train and test set
4    train,test= train_test_split(deu_eng,test_size=0.2,random_state= 12)
```

Fig. 3.2.7.1 Model creation

It's time to encrypt the words. As the input sequences, we will encode Bangla sentences and English sentences as the target sequences. For both the train and the trial datasets, this has to be achieved.

```
1    # prepare training data
2    trainX = encode_sequences(deu_tokenizer, deu_length, train[:, 1])
3    trainY = encode_sequences(eng_tokenizer, eng_length, train[:, 0])
4
5    # prepare validation data
6    testX = encode_sequences(deu_tokenizer, deu_length, test[:, 1])
7    testY = encode_sequences(eng_tokenizer, eng_length, test[:, 0])
```

Fig. 3.2.7.2 Train and text data separation

Now the exciting part is coming!

We'll start off via defining our Seq2Seq model architecture:

For the encoder, we will use an embedding layer and an LSTM layer

For the decoder, we will use every other LSTM layer observed by means of a dense layer
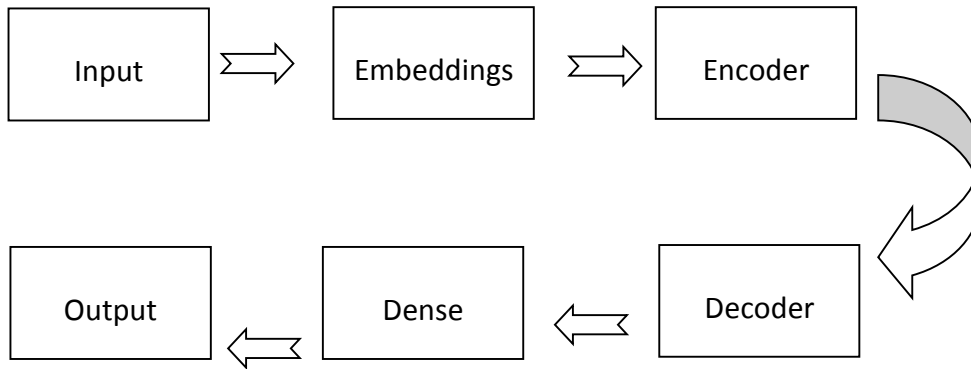
Fig. 3.2.7.3 LSTM layer observation

```
1    # build NMT model
2    def build_model(in_vocab,out_vocab, in_timesteps,out_timesteps,n):
3            model = Sequential()
4            model.add(Embedding(in_vocab, n, input_length=in_timesteps,
5            mask_zero=True))
6            model.add(LSTM(n))
7            model.add(RepeatVector(out_timesteps))
8            model.add(LSTM(n, return_sequences=True))
9            model.add(Dense(out_vocab, activation='softmax'))
10           return model
```

Fig. 3.2.7.4 NMT model

In this design, we use the RMSprop optimizer as it is usually a good choice to work with recurrent neural networks.

```
1    # model compilation (with 512 hidden units)
2    model = build_model(deu_vocab_size, eng_vocab_size, deu_length, eng_length, 512)
3
4    rms = optimizers.RMSprop(lr=0.001)
5    model.compile(optimizer=rms, loss='sparse_categorical_crossentropy')
```

Fig. 3.2.7.5 model compilation

Please be aware that we have used 'sparse_categorical_crossentropy' as the loss function. This is because the feature allows us to use the goal sequence as is,

alternatively of the one-hot encoded format. One-hot encoding the goal sequences using such a massive vocabulary may consume our system's whole memory.

We are all set to start training our model!

We will train it for 30 epochs and with a batch size of 512 with a validation split of 20%. 80% of the data will be used for training the model and the rest for evaluating it. You may change and play around with these hyperparameters.

The ModelCheckpoint() feature will also be used to save the model with the lowest loss of validation. Personally, I prefer this approach to stop early.

```
1   filename = 'model.h1.24_jan_19'
2
3   # set checkpoint
4   checkpoint = ModelCheckpoint(filename, monitor='val_loss',
5                                verbose=1, save_best_only=True,
6                                mode='min')
7
8
9   # train model
10  history = model.fit(trainX, trainY.reshape(trainY.shape[0], trainY.shape[1], 1),
11                      epochs=30, batch_size=512, validation_split = 0.2,
12                      callbacks=[checkpoint], verbose=1)
```

Fig. 3.2.7.6 Saving model

Let's compare the training loss and the validation loss.

```
plt.plot(history.history['loss'])
plt.plot(history.history['val_loss'])
plt.legend(['train','validation'])
plt.show()
```
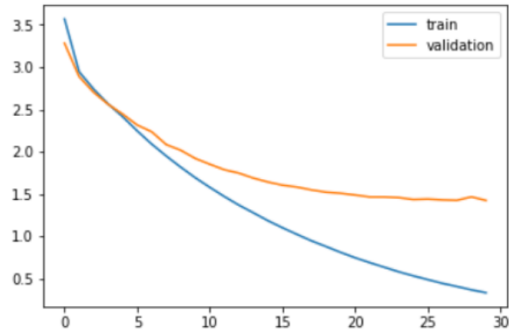

Fig. 3.2.7.7 Train and valid data comparison

As you can see in the above graph, after 20 epochs, the loss of validity stopped decreasing.

Finally, we can load the saved model and make predictions on the unseen data — testX.

```
1   model = load_model('model.h1.24_jan_19')
2
3   preds = model.predict_classes(testX.reshape((testX.shape[0], testX.shape[1])))
```

Fig. 3.2.7.8 Loading trained model

These predictions are sequences of integers. We need to convert these integers to their corresponding words. Let's define a function to do this:

```
1    def get_word(n, tokenizer):
2        for word, index in tokenizer.word_index.items():
3            if index == n:
4                return word
5        return None
```

Fig. 3.2.7.9 Integer to word

Convert predictions into text (English):

```
1    preds_text = []
2    for i in preds:
3        temp = []
4        for j in range(len(i)):
5            t = get_word(i[j], eng_tokenizer)
6            if j > 0:
7                if (t==get_word(i[j-1],eng_tokenizer))or(t== None):
8                    temp.append('')
9                else:
10                   temp.append(t)
11           else:
12               if(t == None):
13                   temp.append('')
14               else:
15                   temp.append(t)
16       preds_text.append(' '.join(temp))
```

Fig. 3.2.7.10 Convert prediction into English

Let's place in the test dataset the original English sentences and the expected sentences:

```
pred_df = pd.DataFrame({'actual' : test[:,0], 'predicted' :
                        preds_text})
```

Fig. 3.2.7.11 Separating predicted and actual data for display

To see how our template works, we can randomly print some real versus expected instances:

```
# print 15 rows randomly
pred_df.sample(15)
```

3.2.7.12 Print result(1)

	actual	predicted
4719	id like to know why	id like to know why
6533	i dont need a car	i dont need a car
8505	youve got a fever	you have a fever
6081	dont come again	dont come again
5615	im tired of boston	im am boston
868	i mean no disrespect	i dont want to be
1415	take a number	take a number
3561	what does it say	what does it
1292	shes still young	she is young
7759	youre out of order	youre not normal
1285	get lost	get away
6900	tom wouldnt blame you	tom wouldnt blame you
4161	youre in love	are love
1427	he noticed i was there	he saw ill right
2987	tom rarely smiled	tom seldom smiled

Fig. 3.2.7.13 Print result(2)

Our Seq2Seq model does a respectable job. But there are several situations the place it misses out on grasp the key words. For example, it translates "I'm worn-out of boston" to "I'm boston" [23].

These are the challenges you will face on an everyday foundation in NLP. But these are not immovable obstacles. We can mitigate such challenges with the aid of the use of greater education information and constructing a higher (or extra complex) model.

3.2.8 End Notes

The results are pretty promising, even with a very simple Seq2Seq model. By using a more sophisticated encoder-decoder template on a larger dataset, we can easily improve this efficiency.

Another experiment I can assume of is trying out the seq2seq approach on a dataset containing longer sentences. The extra you experiment, the more you'll learn about this tremendous and complex space.

3.3 Data Collection Procedure

For the research purpose, we collect Bangla language database. As our system take voice command, so we use google recognizerIntent to get better voice command. Also we train lot of data for our AI system. We take user data through a web server. From which we collect voice command and response of the particular command that user wish to have from the application.

3.4 Implementation Requirements

For implementing our project we use android studio obviously updated version. Project develop in android X version. Server language that we use php and we use MySQL database. As we take voice command, for getting better experience we should give a clear voice command. So we use goggle recognizerIntent to get better voice command. Also we use a virtual device for test run. As this project is android base so we also use android smartphone for implement. But we need at least android version of 4.1 jelly bean.

CHAPTER 4

EXPERIMENTAL RESULTS AND DISCUSSION

4.1 Introduction

To implement a user friendly application that have voice assistant, first we take a look over those applications or devices that has existing voice assistant feature. Most the companies have voice assistants that understands English and other foreign languages. But they have not worked on Bengali language so we have made our aim to build a voice assistant that understands Bengali and English language and performs actions according to user command. We have included other features that will help user to use the application easily. There is a login and sign up feature in our application. To get into the application user needs to login first. This feature makes our application more secure. There are other features that is needed such as news, weather, internet search, to do task and more. All the features are connected with voice assistant. User can have all these features at a same place moreover they can give command to the assistant in Bengali language.

4.2 Experimental Results

In this paper, we have tried to develop an application which is mainly based on voice assistant. Modern world is running towards Artificial Intelligence. We are not different from them. The voice assistants of giant companies influenced us to implement a voice assistant that have various features. We tried to keep ourselves little unique from them because we have trained our voice assistant in Bengali language. Bengali language is the mother tongue of Bangladeshi people. In our application there are other features that makes the easy use of the user. The features that we have included they are login and registration system, splash screen, dashboard, news, weather, to do task, internet search and more.

4.2.1 Splash Screen

Splash screen is a particular screen that we have used as the startup screen. It displays some animation to the user to look at while the application is loading [21]. This animation appears for a specific amount of time with the basic introductory information [22]. In this section we have used android animation, animation library, two text view. This section actually welcomes the user to our application.

4.2.2 Login

This is the process from where user gets the permission to access into our application. In this section users have to enter valid email and password which is registered before. Then the users have to click on the login button to access into the application. In this section we have used two edit text which represents email and password.

4.2.3 Registration

This section is a mandatory process entering into the application. User have to enter their email, password and full name to register. Then the user will click on register. After registering user will be moved to login section. For this, we have used three edit text and two button where one button represents register and another one represents go to login.

4.2.4 Homepage

When a user login to our application, it will be directed the user to the homepage. Inside homepage there are two different sections called Dashboard and AI Voice Assistant. User can give voice command to the voice assistant or user can tap on the features given in the dashboard.

4.2.5 Dashboard

It is a graphical user interface that shows the key performances of our application. User can have various sections in one place. To create this section, we have used android navigation drawer, fragment, API, image library for android, fat button. User can tap on the features that is available in the dashboard and enjoy the features. In the dashboard

section there are various features. It is very helpful for the user and user do not need to use other applications that contains the features available in the dashboard.

4.2.6 AI Voice Assistant

It is based on Artificial Intelligence that performs various actions after hearing the users command. In this section we have used language library of google. This voice assistant can understand Bengali and English language. This assistant is trained successfully in Bengali language. User can give command to the voice assistant in his/her preferred language. This is the unique feature that is available in our voice assistant. When user gives any command the voice assistant hears the command and perform action. We have added some features which works by hearing user voice command. The features are:

4.2.6.1 News

It provides the daily news as it hears the voice command from the user. We have used image view and text view to create this section. User can have his/her preferred news at a same place.

4.2.6.2 Weather

It provides the daily weather report by hearing the voice command from the user. To create this section, we have used weather API. This feature first read the location of the user then it gives the weather report based on user's location.

4.2.6.3 To do Task

It provides user the daily task he has. It can give reminder to the user if he/she has any meeting or something else. User can note down his daily task using the voice assistant.

4.2.6.4 Internet Search

This section can search anything from the internet by hearing command from the user by using voice assistant. This makes search easier. User do not need to type on the keyboard to search anything from the internet. User needs only give command to the voice assistant and the assistant will perform the action.

4.3 Descriptive Analysis

The application AIRI has voice assistant that can perform action according to user command. We tried to test our application if it is working properly or not.

For that reason, we give some people our application to try out the features that we have implemented in our application. The voice assistant can understand two languages Bengali and English. User can give command in his/her preferred language. As we are from Bangladesh, we prefer Bengali language more than English. The users that we gather they tried the voice assistant and they are satisfied with the voice assistant. Though a tradition of artificial intelligence is going higher with the passage of time. User wants something that makes their work easier. They prefer such applications those are easy to use. We tried out find out the requirements that user wants to have in our application. By following the needs of the users, we have developed an application that is user friendly. There are a lots of features available in our application.

The voice assistant can show the daily news, provide the weather detecting user's location, can search anything on internet, provide to do task, can answer some selective questions. We are trying to add more features in our application. Our main concern was Bengali language recognition by the voice assistant and we have successfully implement this feature. When a user turns the language setting in Bengali in his/her phone then the voice assistant understands commands in Bengali. We are trying to improve this feature that user can use more easily. The features that are available in the application will be updated in the future. We have an aim to pass a long term with the help of this application. More features are coming to add on this application. Users are liking our application and they want us to add more features that they need in daily basis. Our voice assistant has a unique feature that is liked by the users mostly. We are

hoping to implement more unique features that will make our application far better. After analyzing the features of the application we came to a decision that our voice assistant is working properly and available features are user friendly.

4.4 Summary

Modern technologies are making the world better day by day. The future of technology is based on Artificial Intelligence and voice assistant is a precious part of modern technology. So for the betterment we tried to implement something that can reflect our mother tongue. There are so many existing voice assistants but no one has trained their voice assistant in Bengali language. As English is international language, they have used English and other foreign languages in their voice assistant. This the reason that we have made our decision to develop an application that will have the voice assistant and the assistant will be trained in Bengali language. We are successful in doing this. Our voice assistant understands commands given in Bengali and English language. It performs action according to user command. Firstly, we have included some common features that is used by the user on a daily basis. These features can be used by tapping or by giving voice command to the assistant. Our application is secure enough for the user. Though there is login system and user can login after registering in the sign up section. By doing registration, user can login to our application. This option keeps user's personal data and information secure. There will be no privacy issue because of the login system. The voice assistant performs action quickly that user can have any feature within a single command. Finally, we have experiment with the existing feature of voice assistant in our own Bengali language. This will be very helpful for Bangladeshi people.

CHAPTER 5
SUMMARY, CONCLUSION AND IMPLICATION FOR FUTURE RESEARCH

5.1 Summary of the Study

In present days, technology develop day by day. Now almost every step we use different machine or software for our comfortable. And Artificial Intelligence is one of the best technology in modern life. It makes our life easier. In our daily needs, along with clothes, food & shelter, a smartphone also becomes the fourth human need. And that's why we try to build an AI software that interact with our daily needs. It will help a user on every step. Like in the morning, it provide daily news along to user choice. Then suggest to-do task, that need to done. Also we provide here voice command assistant that will help you to give any command by voice.

5.2 Conclusions

In this report, we have present an Artificial Intelligence with daily life interaction. Here we provide a system that make our life easier and faster. We don't need to provide command again and again. AI simply detect our daily choice and provide result on user basis. And the big thing is we don't need to give command on only English language here user can give command on Bangla language. For this feature people from our country give command easily. In our system the voice recognize system recognize our Bangla voice 97% accurately. That's a huge plus point for us. We also provide todo task. User can done his daily task properly. User get notification on daily necessary task. We also provide few emergency contact number of our country. Like Fire services, Police Station, Hospital, GOVT emergency dial number etc. All this thing make our system perfect and extraordinary.

5.3 Implication for Further Study

Artificial intelligence (AI) will transform the world later this century. Now a days everyone depend on AI or machine. We face machine on our every step. So to make our life easier we continue develop our AI.

We already add many feature on our system like News along to user choice, Weather prediction, To-do task, Voice command, Useful contact number etc. In future we add more feature that can improve our daily needs. Like making a call by voice command, flash light on by voice command, more accurate voice reorganization assistant etc.

REFERENCES

[1] Ahmed Sallam, available at <<https://sallamah.weebly.com/ai-2015-f.ht>> last access on 29-10-19 at 12:49 AM

[2] Life and Motivation, available at <<https://lifeandmotivation.blogspot.com/>> last access on 29-10-19 at 12:00 AM

[3] Tech.Co, available at <<https://tech.co/news/6-main-ways-technol>> last access on 29-10-19 at 12:50 AM

[4] Jeffbullas, available at <<https://www.jeffbullas.com/artificial-in>> last access on 29-10-19 at 12:52 AM

[5] Course Hero, available at << https://www.coursehero.com/file/p1np5vv/>> last access on 29-10-19 at 12:57 AM

[6] wikipedia, available at <<https://en.wikipedia.org/wiki/Google_Assistant>> last access on 29-10-19 at 12:59 AM

[7] wikipedia, available at <<https://en.wikipedia.org/wiki/Siri>> last access on 29-10-19 at 1:02 AM

[8] linkedin, available at <<https://in.linkedin.com/in/capturingeye>> last access on 29-10-19 at 1:04 AM

[9] wikipedia, available at <<https://en.wikipedia.org/wiki/Amazon_Alexa>> last access on 29-10-19 at 1:08 AM

[10] Deccan Chronicle, available at <<https://www.deccanchronicle.com/technolo>> last access on 29-10-19 at 1:09 AM

[11] wikipedia, available at <<https://en.wikipedia.org/wiki/Amazon_Alexa>> last access on 29-10-19 at 1:11 AM

[12] wikipedia, available at <<https://en.wikipedia.org/wiki/Bixby_(virtual_assistant)>> last access on 29-10-19 at 1:15 AM

[13] microsoft, available at <<https://answers.microsoft.com/en-us/windows/forum/windows_10-desktop/virtual-assistant/ab354fad-4c36-473f-9f0c-b07ecf4f2677>> last access on 29-10-19 at 1:15 AM

[14] fandom, available at <<https://winodws.fandom.com/wiki/Cortana>> last access on 29-10-19 at 1:17 AM

[15] jamf|Now, available at <<https://www.windowscentral.com/history-cortana-microsofts-digital-assistant>> last access on 29-10-19 at 1:30 AM

[16] kslib, available at <<https://kslib.info/DocumentCenter/View/6669/2017AI?bidId=>> last access on 29-10-19 at 1:33 AM

[17] analytics vidhya, available at <<https://www.analyticsvidhya.com/blog/2019/06/understanding-transformers-nlp-state-of-the-art-models/>> last access on 29-10-19 at 1:35 AM

[18] ground ai, available at <<https://www.groundai.com/project/mass-masked-sequence-to-sequence-pre-training-for-language-generation/1>> last access on 29-10-19 at 1:37 AM

[19] processing, available at <<https://processing.org/tutorials/text/>> last access on 29-10-19 at 1:40 AM

[20] coursera, available at <<https://www.coursera.org/lecture/natural-language-processing-tensorflow/notebook-for-lesson-2-Sydkf>> last access on 29-10-19 at 1:45 AM

[21] smallenvelop, available at <<https://smallenvelop.com/display-loading-icon-page-loads-completely/>> last access on 29-10-19 at 1:47 AM

[22] geeksforgeeks, available at <<https://www.geeksforgeeks.org/android-creating-a-splash-screen/>> last access on 29-10-19 at 1:49 AM

[23] medium, available at <<https://medium.com/analytics-vidhya/a-must-read-nlp-tutorial-on-neural-machine-translation-the-technique-powering-google-translate-c5c8d97d7587>> last access on 29-10-19 at 1:50 AM

[24] calvium, available at <<https://calvium.com/voice-technology-challenges-and-opportunities/>> last access on 29-10-19 at 1:52 AM

APPENDIX

Appendix A: Research Reflection

The purpose of this Appendix is to produce associate degree introduction to analysis meal. The cluster research was a difficult and pleasant expertise typical of the course as an entire. We have had very little exposer to cluster work university. So, it absolutely was a pleasant modification to be a part of a good and dynamic team.

The experience taught us that planning and crafting responses takes a longer time in teams than on your own. The intensive effort needed was ultimately an honest issue. In our cluster, we have a tendency to area unit perpetually developing and refinement one another's ideas. We input data from many user. This analysis results would facilitate them to their future cultivation by all suggests that.

Appendix B: Related Issues

Collecting Bangla and English database is quite difficult. Because Bangla resource is quite little in the web. So we input thousands of data by user to our server.

AIRI

10	Submitted to Great Oak High School Student Paper	<1%
11	Submitted to Universiti Utara Malaysia Student Paper	<1%
12	www.windowscentral.com Internet Source	<1%
13	Submitted to Marshall High School Student Paper	<1%
14	smoorelaws30012011.weebly.com Internet Source	<1%
15	Submitted to CSU, San Jose State University Student Paper	<1%
16	Submitted to Wiltshire College Student Paper	<1%
17	Submitted to Management Development Institute Of Singapore Student Paper	<1%
18	Submitted to Arab Open University Student Paper	<1%
19	Submitted to KDU College Sdn Bhd Student Paper	<1%
20	Submitted to The College of New Jersey Student Paper	<1%

Submitted to University of Ulster

21 Student Paper <1%

22 **Submitted to Sunway Education Group**
Student Paper <1%

23 **Submitted to Baylor University**
Student Paper <1%

24 en.m.wikipedia.org
Internet Source <1%

25 ahtresidential.com
Internet Source <1%

26 kancy.com
Internet Source <1%

27 **Submitted to SAE Institute, London**
Student Paper <1%

28 **Submitted to Manukau Institute of Technology**
Student Paper <1%

29 www.ivoryresearch.com
Internet Source <1%

30 export.arxiv.org
Internet Source <1%

31 **Submitted to Griffith College Dublin**
Student Paper <1%

32 **Submitted to London Metropolitan University**
Student Paper <1%

33	answers.themler.io Internet Source	<1%
34	Submitted to Korea University Student Paper	<1%
35	scholar.uwindsor.ca Internet Source	<1%
36	Submitted to City University Student Paper	<1%
37	Submitted to Sri Lanka Institute of Information Technology Student Paper	<1%
38	Submitted to University of Westminster Student Paper	<1%
39	Hua Zhong, Lingming Zhang, Sarfraz Khurshid. "TestSage: Regression Test Selection for Large-Scale Web Service Testing", 2019 12th IEEE Conference on Software Testing, Validation and Verification (ICST), 2019 Publication	<1%
40	www.reference.com Internet Source	<1%
41	Submitted to TAR University College Student Paper	<1%
42	Submitted to Ivy Tech Community College	