Towards a system to aid communication with Deaf

(Ava)

Submitted in partial fulfillment of

the requirements for the award of the degree of

Bachelor of Information Systems in

Computer Science

Submitted by

Mehreen Najm

Ahmad Zia Yosufi

Under the supervision of

Asst.Prof Rafiullah Momand



Information Systems Department

Computer Science Faculty

Kabul University

Kabul, Afghanistan

Spring 2019

**Abstract**

Towards a system to aid communication with Deaf (Ava) is an experimental system that aims to aid communication between a deaf person and an ordinary person by translating the ordinary person’s speech to sign language Using Sequence to Sequence Neural Machine Translation. A speech recognizer recognizes speech from the ordinary person and the system then synthesizes the appropriate sequence of signs in American Sign language (ASL) using a specially developed avatar. By using a phrase lookup approach to language translation, which is appropriate for the highly constrained discourse in society, we were able to build a working system that we could evaluate. We used a model where English language is translated into American Sign Language Moreover, we used SGML for displaying the output of translated ASL using avatar.

**Contents**

1. **Chapter 1 (Introduction)**

1.1 Problem statement………..………….……………………………………………….1

1. **Chapter 2 (State of Research)**
   1. Background and Research……………………………………………….…………….2

2.1.1 Literature Review...........................................................................................…...4

* 1. Motivation………..……………………………………………………..…….……….8
  2. Objectives……………………………………………………………….…………….8

1. **Chapter 3 (Work Done)**
   1. State of Research………….……………………………………………………….…..9

**4. Chapter 4 (Result and Discussion)**

4.1 Preparing the Dataset and the Module for train………………………………..……….….13

4.1.1 Dataset …………………………………………………………………….13

4.2.2 Implementation…………………………………………………………….14

4.2 Result……………………………………………………………………….…..……….….31

**5. Chapter 5 (Conclusion)**

5.1 Future Works……………………………………………………………………………......32

**6. Acknowledgments………………………………………………34**

**7. Bibliography………………………………….…………………35**

**8. Appendix A…………………………..……….…………………36**

**9. Appendix B………………………………….…………………41**

**List of Figures**

**2.1 Sign representation………………………………….…………………3**

**3.1 Encoder and Decoder illustration…………………………………….10**

**3.2 Seq2Seq model input and output……………………………………..11**

**4.1 Encoder and Decoder formula………………………………………..15**

**4.2 Dataset sample…………………………………………………………16**

**4.3 Dataset sample…………………………………………………………17**

**4.4 Dataset sample…………………………………………………………18**

**4.5 Dataset sample…………………………………………………………19**

**4.6 Dataset sample………………………………………………………....20**

**4.7 Dataset sample…………………………………………………………21**

**4.8 Preparing the Model…………………………………………………..22**

**4.9 preparing the model for train using TenserFlow……………………23**

**4.10 Preparing Dataset for train………………………………………….24**

**4.11 Reading files from directory…………………………………………25**

**4.12 Perform normalization on dataset…………………………………..25**

**4.13 Perform normalization on dataset…………………………………..26**

**4.14 RNN encoder and decoder for Se2Seq model……………………...26**

**4.15 Encoder function……………………………………………………..27**

**4.16 Encoder function…………………………………………………….27**

**4.17 Decoder function…………………………………………………….28**

**4.18 Define Optimizer and loss function………………………………...28**

**4.19 Train the model……………………………………………………...29**

**4.20 calculating the total loss……………………………………………..29**

**4.21 evaluating the model…………………………………………………30**

**4.22 Test the model………………………………………………………..3**

**Chapter 1**

**Introduction**

**1.1 Problem Definition:**

In our modern information and communication society, daily life would be unimaginable without technology [1]. Information and Communications Technology (ICT) is also very useful for people with special needs.

Deaf are people who can’t talk and hear, hearing people are unfamiliar with Deaf because they don’t know their language. They think that if you cannot hear, you can easily access any necessary information by simply reading it in written form, and if you really need to communicate, you can always write your message down. [6]

But this system is not an efficient way to communicate, Deaf have their own language of communication (sign language), they can only understand this language and also they write and read Faster on sign writing. According to survey in 2015, there are 121 “Deaf sign languages” in the world, but there is not such an efficient mechanism where Deaf can easily understand people thoughts. [7]

* 1. **Background and Recent Research**

People who are Deaf have little or no hearing ability. The word “Deaf”, (often written with a capital “D”) typically refers to people who use Sign Language as a primary language – either directly with others who sign or indirectly through an interpreter with people who do not sign.

Sign language for deaf people has special features that are quite different from those of spoken language. Sign language is an iconic language compared to spoken language, which is more of an arbitrary one. Another crucial difference between the two languages is that sign language does not have its own writing system. Therefore, in order to write descriptions of signs, line drawings, photographs and illustrations have commonly been used, but these represent only a very small moment in the process of actual signing. [1]

They may hear environmental sounds, and may even understand some speech, but they identify with what is known as the “Deaf Culture” [2]. There has recently been considerable research activity in developing automatic systems which can understand and output speech to provide information services or to perform transactions with customers [3]. We have been developing a system which enable people to communicate with Deaf. It is an interactive translation system to assist in the completion of a conversation between an ordinary person and a deaf. The system translates the ordinary person speech into American Sign Language (ASL) and displays the signs using a specially-developed avatar. A comprehensive approach to the task of enabling humans who cannot sign to communicate using sign-language would clearly require the development of a general purpose speech to sign language converter. This in turn requires the solution of the following problems:

1. Automatic speech to text conversion (speech Recognition).
2. Automatic translation of English text into a suitable representation of sign language
3. Display of this representation as a sequence of Signs using computer graphics techniques.

**Sign Language Representations:** [1]

A notation system for sign language is strongly needed to advance the study of its structure**,**

Two pioneers, William C. Stokoe, who proposed a notation system for American Sign Language (ASL), and Lynn Friedman, who analyzed ASL from a phonological point of view, worked towards this end. However, their notation systems are rather impractical for general users because they are too technical, much like phonetic alphabets in spoken languages. Therefore, it is necessary to employ another writing system to describe a sign or a signed sentence for everyday purposes. A more suitable sign writing method for this purpose is “SignWriting” by the American movement analyst Valerie Sutton. Her system was applied to Japanese Sign Language (JSL) and the results indicated that the system was effective enough to “write” Japanese Sign Language.

Indeed, there are newsletters and books written in SignWriting in some countries and the practical usage of the system in literacy education has had a great impact on the education of deaf children.

Sutton’s Center for Sutton Movement Writing, Inc. has created the following writing systems.

**1. Sutton Movement Writing**

Sutton Movement Writing is the International Movement Writing Alphabet (IMWA) used to record all human and animal gestures. The IMWA records the details of movement-based languages. The IMWA has been specifically designed for the following five fields:

a. SignWriting: for writing the movements of sign languages

b. DanceWriting: for writing dance choreography

c. MimeWriting: for writing classic pantomime

d. SportsWriting: for writing ice skating and gymnastics routines

e. ScienceWriting: for gesture-based research.

**2. Sign-Symbol-Sequence (formal sign writing):**

The sign symbol sequence (FSW) is the official sequence of symbols in Sutton movement writing. The FSW is used in computer to sort and look up movements in Movement Writing.

1. **Sutton’s SymbolBank**

Sutton’s SymbolBank Database is the official source for all symbols in Sutton Movement Writing.

Sutton’s writing system of sign language is composed with symbols for head and face cues, hand shape, Movement (direction and manner), qualification and punctuation. Signed words written in SignWriting are created by compounding symbols of a highly pictorial design. This system has the following advantages:

1. It is aimed at enabling us to write every sign or signed sentence of every country;

2. Because the Writing is pictorial, natural shapes and movements of signs can be realistically shown;

3. Face expressions and body movements can be depicted, too.



**Figure 2.1 Sign representation of Deaf**

**1.3 Motivation**

Most of the systems that are build are based on a specific place or a specific language for Deaf. Despite TESSA, ATLAS, Scribe4Me and some systems that I mentioned above the problem of communication with deaf still remains and open discussion among researchers therefore, my approach is to suggest an efficient and effective way that help ordinary people to establish a communication link with deaf, this can be possible with international sign language, as international sign language contains signs and movements that are common among Deaf culture. Using this method it is possible to transform any language into sign language.

The second method that is missing from others and motivated me to work on it is making sign per movement instead of each word. In other systems like Braille [9] it used glove and sensors for each movement that is time consuming and make the dictionary large therefore, by making signs per each movement we just need around 600 sign representation which is cost-effective and time saving.

By making such system deaf can easily understand people thoughts and they feel there is a place for them in technology and technology do care about them.

**1.4 Methodology**

We divided our Methodologies in 2 parts:

1. **Research methodology:** Firstly we used literature review for gathering information and had a brief study about the systems which have already exists for Deaf, secondly we compared the existing systems and our suggested solution and found some techniques for better performance and sustainability of our suggested system.
2. **Development Methodology:** We used agile methodology as our datasets are in progress, we used some features and technologies like (TenserFlow,) as well as we have found some algorithms and models for machine translation such us RNN, RNN-encoder, RNN decoder, Seq2Seq model.

**1.1.1 Literature Review:**

This chapter is based on research that we have done to solve the problems such as differences between this system and other similar systems already been built. In this system, we tried our best to overcome challenges we observed in current systems. So according to our researches and findings, there were not any system to convert text/speech or ordinary people into sign language and represent it using an avatar. Briefly, there are systems for Deaf like Tessa but they are limited to a Post Office and they used British Sign language (BSL) instead of American Sign Language (ASL).

**Area 1: TESSA**

TESSA is an experimental system that aims to aid transactions between a deaf person and a clerk in a Post Office. These systems are designed to provide translation of conversational speech between languages with a potentially very large vocabulary. It is an interactive translation system but it operates in a very restricted domain and is designed to assist in the completion of a transaction between a Post Office (PO) clerk and a deaf customer. The system translates the clerk's speech into British Sign Language (BSL) and displays the signs using a specially-developed avatar. [8]

In addition, concatenation of signing is more fluent and controlled for avatar than for video signing, as the exact positioning of the avatar can be manipulated. For these reasons, we decided to display the signs using an avatar, TESSA. Research into methods for capturing signing movements directly from video has been reported.

The alternative is to capture signs using separate sensors for the hands, body and face. This technique appears to capture sufficient movement to generate true and realistic signing from a virtual human. The motion is captured as follows:

1. Cyber gloves with 18 resistive elements for each hand are used to record finger and thumb positions relative to the hand itself.

2. Polhemus magnetic sensors record the wrist, upper arm, head and upper torso positions in three-dimensional space relative to a magnetic field source.

3. Facial movements are captured using a helmet mounted camera with infra-red filters and surrounded by infra-red light emitting diodes to illuminate Scotch light reflectors stuck onto the face. Typically 18 reflectors are placed in regions of interest such as the mouth and eyebrows.

**Area 2: HandTalk**

Founded in 2012, Hand Talk performs digital and automatic translation into the Brazilian Sign Language. HandTalk, which is internationally awarded and a reference in the segment, is run by a friendly virtual interpreter, Hugo, a 3D character that makes communication interactive and easy to understand. For more visit here: <https://www.handtalk.me/sobre>

**Area 3: Braille**

This system includes a smart glove that translates the Braille alphabet, which is used almost universally by the literate deaf blind population, into text and vice versa, and communicates the message via SMS to a remote contact. It enables user to convey simple messages by capacitive touch sensors as input sensors placed on the palmer side of the glove and converted to text by the PC/mobile phone. The wearer can perceive and interpret incoming messages by tactile feedback patterns of mini vibrational motors on the dorsal side of the glove. The successful implementation of real-time two- way translation between English and Braille, and communication of the wearable device with a mobile phone/PC opens up new opportunities of information exchange which were hitherto un-available to deaf blind individuals, such as remote communication, as well as parallel one-to many broadcast. The glove also makes communicating with laypersons without knowledge of Braille possible, without the need for trained interpreters. [9]

**Area 4: Deaf-Mute Communication Interpreter**

The project aims to facilitate people by means of a glove based deaf-mute communication interpreter system. The glove is internally equipped with five flex sensors, tactile sensors and accelerometer. For each specific gesture, the flex sensor produces a proportional change in resistance and accelerometer measures the orientation of hand. The processing of these hand gestures is in Arduino. The glove includes two modes of operation –training mode to benefit every user and an operational mode. The concatenation of letters to form words is also done in Arduino. In addition, the system also includes a text to speech conversion (TTS) block which translates the matched gesturesi.e. Text to voice output. [10]

**Area 5: Arabic Sign Language Interpreter**

The real difficulties arise when a deaf person wants to communicate with a non-deaf person. Usually both will get frustrated in a very short time. For this reason, there have been several attempts to design smart devices that can work as interpreters between the deaf people and others. These devices are categorized as human-computer-interaction (HCI) systems [11].

The adaptation of mobile devices makes sign language translation more attractive and more valuable. Using mobile devices instead of PC-base in sign languages presents several advantages. With PC base, term of anywhere and any place in deaf learning cannot be applied. With mobile devices one can obtain communication more realistic, a wider usage of sign language applications becomes possible and practical.

In recent years, several research projects in developing sign language animations system have been developed [12]. Some previous projects have made efforts in translating English text into Sign Language Animation, but none have proposed practical systems for translating Arabic text into Arabic Sign Language use mobile technologies and mobile devices. And also, most of the previous systems are PC-base. The adoption of mobile devices in developing sign language animation systems is motivated by several considerations: they help deaf to upgrade quality of human-human communication by evolving animations; they have a positive impact on factors such as human-mobility and likeability; they can have a positive effect on a deaf perception of deaf learning experience because they can attract deaf attention. [13]

**Area 6: Scribe4Me**

Sound plays an important role in communication and contextual awareness about interesting events and information. For this reason mobile support for better sound information awareness would be of great value to the deaf and hard-of-hearing. This system present the design and evaluation of a mobile sound transcription tool for the deaf and hard-of-hearing called Scribe4Me. When a user presses a button on her Scribe4Me PDA, the last 30 seconds of sound is uploaded, transcribed and sent back to her as a text message. Transcriptions include dialog and descriptions of environ-mental sounds. Scribe4Me is unique in providing support for both speech and text. [16]

**Area 7: A Spanish speech to sign language translation system for assisting Deaf-Mute People**

In Spain, during the last 20 years, there have been several proposals for normalizing Spanish Sign Language, but none of them has been accepted by the deaf-mute people community. From their point of view, these proposals tend to constrain the sign language, limiting its flexibility. In 1991, MA. Rodríguez [14] carried out a detailed analysis of Spanish Sign Language (SSL). She showed the differences between the sign language used by deaf-mute people and the standardization proposals.

This system translates officer explanations into sign language for deaf-mute people. The translation system is composed by a speech recognizer (for decoding the spoken utterance into a word sequence), a natural language translator (for converting a word sequence into a sequence of gestures belonging to the sign language), and a 3D avatar animation module. [15]

**Area 8: ATLAS**

Sign Languages (SLs) are visual languages used by deaf people to convey meaning. SLs rely on signs as lexical units instead of words used in common languages. Italian deaf people resort to Italian Sign Language (LIS) within their communities and it can be considered as the main way of communication of 60.000 Italian deaf individuals [17].

A system for the computer-assisted translation from Italian Language to LIS that provides the output of the translation resorting to a virtual avatar. [19]

The ATLAS system is designed to get written text as input and to perform the translation, resorting to two different translators: a statistical one and a rule based one. The statistical translator is based on MOSES [18], an open source statistical translator that automatically trains the translation models for any language pair. The Rule Based translator is based on a traditional rule-based approach. The input sentences are interpreted in terms of an ontology-based logical representation, which acts as input to a linguistic generator that produces the corresponding LIS (Italian Sign Language) sentence.

**1.4 Objectives:**

We will suggest an efficient mechanism/system that Deaf can easily communicate with ordinary people to understand their thoughts. As Deaf is a part of society and they need to make a connection with people around them. Our system gets the text/speech from smart phone, convert it to sign language using Sequence to Sequence Neural Machine translation and represent the speech using 3D character.

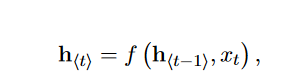
**Chapter 2**

* 1. **State of Research**

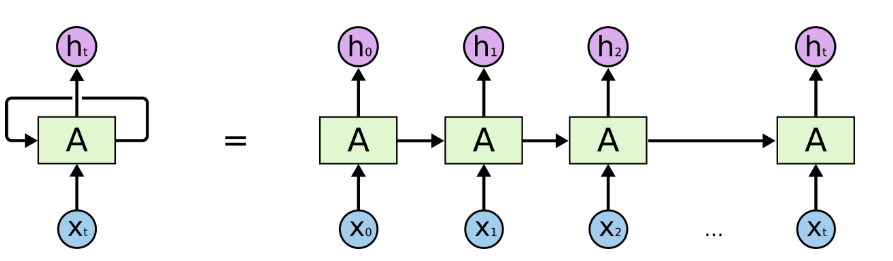
For many people who have been profoundly deaf from a young age, signing is their first language so they learn to read and write English as a second language [4]. As a result, many deaf people have below-average reading abilities for English text and prefer to communicate using sign language [5].

In this chapter we explained what we have used for our development methodology:

* 1. **RNN Algorithm:** Recurrent Neural Networks is a neural net-work that consists of a hidden state hand an optional output which operates on a variable-length sequence= (x1,... xT). At each timesheet, the hidden stateh〈t〉of the RNN is updated by

****

Each member of the output is produced using the same **update rule** applied to the previous outputs.

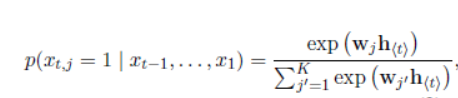


 the current hidden state **h(t)**is a function **f** of the previous hidden state **h(t-1)** and the current input **x(t)**. The **theta** are the parameters of the function f. The network typically learns to use h(t) as a kind of lossy summary of the task-relevant aspects of the past sequence of inputs up to t.

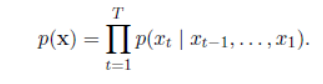
Where is a non-linear activation function may be as simple as an element-wise logistic sigmoid function and as complex as a long short-term memory (LSTM) unit. [20]

An RNN can learn a probability distribution over a sequence by being trained to predict the next symbol in a sequence. In that case, the output at each time step t is the conditional distribution p (xt | xt-1… x1). For example, a multinomial distribution (1-of-K coding) can be output using a softmax activation function. [21]

For all possible symbols j = 1 ,….,K, where wj are the rows of a weight matrix W. By combining these probabilities.

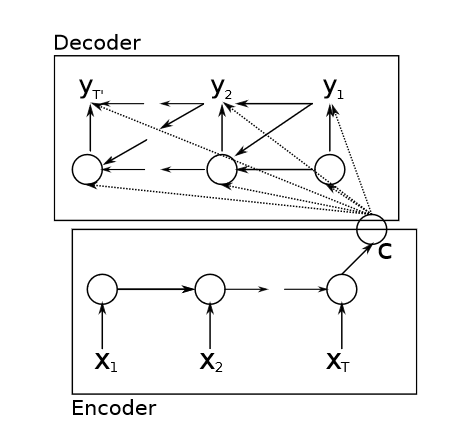
****

we can compute the probability of the sequence x using: [22]

****

From this learned distribution, it is straightforward to sample a new sequence by iteratively sampling a symbol at each time step.

* 1. **RNN Encoder–Decoder:** neural network has an architecture that learns to encode a variable-length sequence into a fixed-length vector representation and to decode a given fixed-length vector representation back into a variable-length sequence. From a probabilistic perspective, this new model is a general method to learn the conditional distribution over a variable-length sequence conditioned on yet another variable-length sequence, e.g. p(y1,...,yT′|x1,...,xT), where one should note that the input and output sequence lengths T and T′ may differ. The encoder is an RNN that reads each symbol of an input sequence x sequentially. As it reads each symbol, the hidden state of the RNN changes according to Figure (1). After reading the end of the sequence (marked by an end-of-sequence symbol), the hidden state of the RNN is a summary c of the whole input sequence.



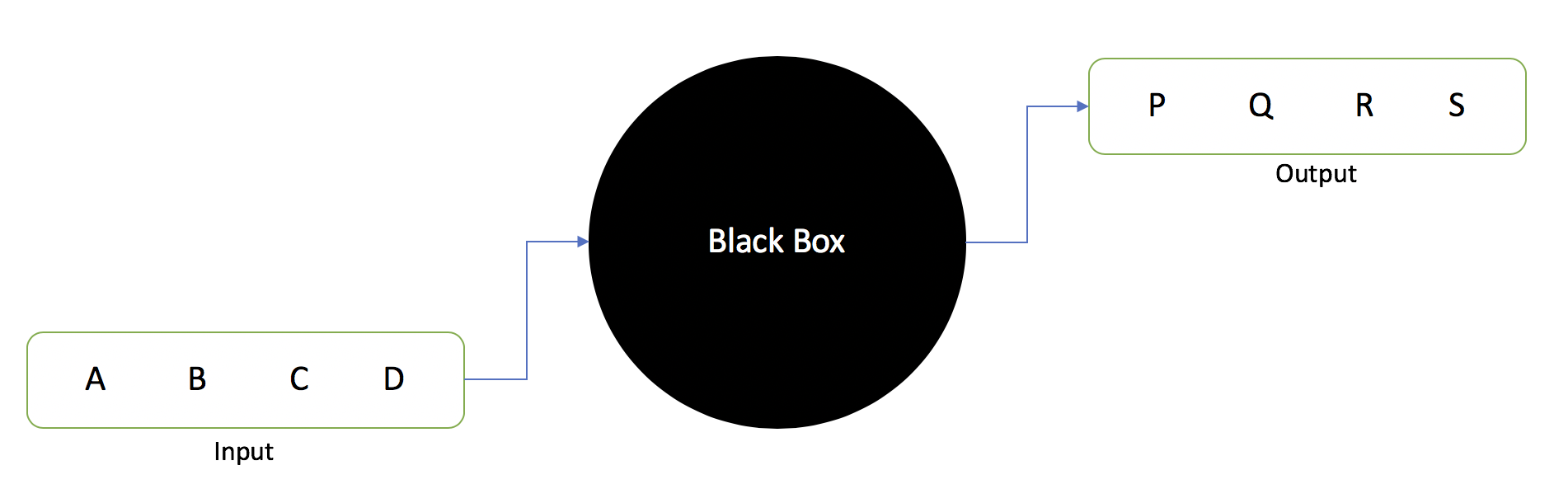
**Figure 3.1 Encoder and Decoder illustration**

The decoder of the proposed model is another RNN which is trained to generate the output sequence by predicting the next symbol *yt* given the hidden state *h<t>*.



**2.3 Seq2Seq Model:** Sequence-to-sequence models have achieved state-of-the-art results on many natural language processing problems including automatic speech recognition, neural machine, conversational modeling and many more. These models learn to generate a variable-length sequence of tokens (e.g. texts) from a variable-length sequence of input data (e.g. speech or the same texts in another language). With a sufficiently large labeled dataset, vanillaSeq2Seq can model sequential mapping well, but it is often augmented with a language model to further improve the fluency of the generated text. Because language models can be trained from abundantly available unsupervised text corpora which can have as many as one billion tokens, leveraging the rich linguistic in-formation of the label domain can considerably improveSeq2Seq’s performance. [23]

A basic Seq2Seq model comprises an encoder that maps an input sequence x= (x1,..., xT)into an intermediate representation, and a decoder that in turn generates an out-put sequence= (y1,...,yK)from h. The decoder can also attend to a certain part of the encoder states with an attention mechanism. The attention mechanism is called hybrid attention, if it uses both the content and the previous context to compute the next context. It is soft if it computes the expectation over the encoder as opposed to selecting a slice out of the encoder states. For the automatic speech recognition (ASR) task, theSeq2Seq model is called an acoustic model (AM) and maps a sequence of spectrogram features extracted from a speech signal to characters.



**Figure 3.2 Seq2Seq model input and output**

**2.4 TenserFlow:** Machine translation using deep neural networks achieved great success with sequence-to-sequence models that used recur-rent neural networks (RNNs) with LSTM cells. [24]

The basic sequence-to-sequence architecture is composed of an RNN encoder which reads the source sentence one token at a time and transforms it into a fixed-sized state vector. This is followed by an RNN decoder, which generates the target sentence, one token at a time, from the state vector. While a pure sequence-to-sequence recurrent neural network can already obtain good translation results, it suffers from the fact that the whole input sentence needs to be encoded into a single fixed-size vector. This clearly manifests itself in the degradation of translation quality on longer sentences and was partially overcome by using a neural model of attention. [24]

Tensor2Tensor (T2T) is a library of deep learning models and datasets designed to make deep learning research faster and more accessible. T2T uses Tensorflow [25] through-out and there is a strong focus on performance as well as usability. Through its use of Tensor-Flow and various T2T-specific abstractions, researchers can train models on CPU, GPU (single or multiple), and TPU, locally and in the cloud, usually with no or minimal device-specific code or configuration. Development began focused on neural machine translation and so Tensor2Tensor includes many of the most successful NMT models and standard datasets. It has since added support for other task types as well across multiple media (text, images, video, and audio).

**2.5 Unity Game engine:**

Unity is a [cross-platform](https://en.wikipedia.org/wiki/Cross-platform) [game engine](https://en.wikipedia.org/wiki/Game_engine) developed by [Unity Technologies](https://en.wikipedia.org/wiki/Unity_Technologies), first announced and released in June 2005 at [Apple Inc.](https://en.wikipedia.org/wiki/Apple_Inc.)'s [Worldwide Developers Conference](https://en.wikipedia.org/wiki/Apple_Worldwide_Developers_Conference) as a [Mac OS X](https://en.wikipedia.org/wiki/MacOS)-exclusive game engine. As of 2018, the engine had been extended to support more than 25 platforms. The engine can be used to create [three-dimensional](https://en.wikipedia.org/wiki/Three-dimensional_space), [two-dimensional](https://en.wikipedia.org/wiki/Two-dimensional_space), virtual reality, and augmented reality games, as well as [simulations](https://en.wikipedia.org/wiki/Computer_simulation) and other experiences [26].Unity is a cross-platform engine [27].

Unity is a game engine that tries to democratize 3D and 2D game development. It is an extremely powerful tool, which allows you to create complex games with cutting edge graphics for many platforms (from PC to mobile to consoles). In the last years, the capabilities of the engine grew and grew to a point where it is hard to keep track. But for now, you don’t have to worry about that. When using ML-Agents we can ignore many of those features without feeling bad. [28]

**Unity Features**

Unity 3D comes loaded with a ton of professional tools for both programmers and artists. Unity provides a workspace that combines artist-friendly tools with a component-driven design that makes game development pretty darn intuitive. Both 2D and 3D development is possible in Unity, with 2D physics handled by the popular [Box2D engine](https://box2d.org/). Unity uses a component-based approach to game dev revolving around [prefabs](https://docs.unity3d.com/Manual/Prefabs.html). With prefabs, game designers can build objects and environments more efficiently and scale faster. With powerful shades, physics-based materials, post-processing, and high-resolution lighting systems, Unity can deliver impressive graphics across the board.

**2.6 SGML:**

Structured documentation is built upon structured elements: chapters, sections, paragraphs, etcetera, where all elements are clearly labeled for what they are: references, program output, etc. No explicit information about how the document should be rendered is given; only about its structure (and content). When there are explicit rules for presentation, they are kept outside the SGML-document.

This allows for automatic processing of the documents, without waiting for AI systems. It encourages authors to concentrate on structure, which conveys meaning.

Standard Generalized Markup Language is a standardized language intended to facilitate the authoring of structured documentation. More specifically, it is a meta-language. You never actually type SGML, but SGML is used to describe a document type specific structured language (this is called a DTD, a Document Type Definition), which defines how specific documents might be structured (written).

SGML is a markup language. All SGML documents include text, mixed with *tags*, which delimits *elements*. SGML allows several syntaxes to be used, but we'll stick with the reference syntax, the most common, where tags are enclosed between angle brackets, < and >. Here is an example:

<article>

<title>The Foo software</title>

<para>

Foo is very fast. And its documentation can be read easily.

</para>

If it looks like HTML to you, it is because HTML is (theoretically) a DTD of SGML.

Elements have a *content*. For instance, the content of the above para element is "Foo is very fast. And its documentation can be read easily.

Elements can have *attributes* to indicate more information. For instance:

<example tested="true">

\*c++;

</example>

SGML provides an internationally recognized, *non-proprietary* language for writing your own markup schemes. Markup is the text that is added to the data of your files in order to convey particular info nation about that data. In a word processor, the markup is the proprietary codes that the software inserts into your text files to indicate which words should be printed in a certain font, which paragraphs should be centered, where page breaks occur etc. In a database system, the markup is the proprietary codes which indicate where one field or record ends and another begins, and so on.

**Chapter 3**

**Work Done**

**3.1 Preparing the Dataset and the Module for train**

**3.1.1 Dataset:**

The dataset used is a sample of common words and sentences that deaf are using daily.

The font that we used is American Sign Language which is published on 2014.

We have searched many sources for gathering dataset such as Wikipedia pages, some books like (Goldilocks & the Three Bears in American Sign Language, The Book of Jonah) , Signbank and other sources, but the dataset that was used for this project is a subset of a much larger dataset, as described in <http://www.signbank.org/signbank.html> . As dataset in this website was so diffused so we gathered and sorted dataset first. In this website there are more than 10000 words in ASL and some words with frequencies.

Our dataset has the following feature vectors:

𝠀񀀁񀀉񈗥񈗵񋸦𝠃𝤝𝤨񀀁𝤎𝣤񀀉𝣰𝣮񈗥𝤏𝤇񈗵𝣱𝤑񋸦𝤆𝤜 Come

𝠀񀀁񆇡񈗥񋾡𝠃𝤘𝤰񋾡𝣴𝣵񀀁𝣹𝤒񈗥𝣣𝤇񆇡𝤋𝤘 Disappointed

𝠀񀀁񆕁񇆥񆡁񆡁񍘡𝠃𝤧𝤺񆡁𝤜𝤲񍘡𝣴𝣵񆡁𝤐𝤲񆕁𝤔𝤔񇆥𝤎𝤢񀀁𝣹𝤑 Red

𝠀񀀁񆨡񋲡񍦁𝠃𝤟𝤬񍦁𝣴𝣵񀀁𝣾𝤎񆨡𝤓𝤚񋲡𝤓𝤤 Throw in your mouth

𝠀񀀙񈁑񋾡𝠃𝤘𝤵񀀙𝣝𝣲񋾡𝣴𝣵񈁑𝣟𝤗 He

𝠀񀀙񈁑񋾡𝠃𝤘𝤵񀀙𝣝𝣲񋾡𝣴𝣵񈁑𝣟𝤗 She

𝠀񀀡񀀁񆫡𝠃𝤛𝤚񀀁𝤄𝣼񆫡𝤉𝣲񀀡𝣱𝣼 111

𝠀񀀡񂤁񁳑񈙇񇆥𝠃𝤯𝤡񀀡𝣝𝣬񂤁𝣱𝣶񁳑𝤆𝣫񈙇𝤑𝤇񇆥𝤈𝤓 155

𝠀񀀡񂤁񁳑񈙇񇆥𝠃𝤯𝤡񀀡𝣝𝣬񂤁𝣱𝣶񁳑𝤆𝣫񈙇𝤑𝤇񇆥𝤈𝤓 One hundred fifty five

𝠀񀀡񂤁񁳑񈙇񇆥𝠃𝤯𝤡񀀡𝣝𝣬񂤁𝣱𝣶񁳑𝤆𝣫񈙇𝤑𝤇񇆥𝤈𝤓 One-hundred-fifty-five

𝠀񀀡񂤁񃋁񀭡𝠃𝤰𝤖񀀡𝣜𝣷񂤁𝣰𝤁񃋁𝤅𝣹񀭡𝤙𝣷 163

Till now we gathered 1500 small sentences with their sign representation informal sign writing.

**4.2.2 Implementation:**

As we mentioned above a comprehensive approach to the task of enabling humans who cannot sign to communicate using sign-language would clearly require the development of a general purpose speech to sign language converter. This in turn requires the solution of the following problems:

1. Automatic speech to text conversion (speech Recognition).
2. Automatic translation of English text into a suitable representation of sign language.
3. Display of this representation as a sequence of Signs using computer graphics techniques.

As we mentioned, the most suitable sign representation is “international sign writing”, we try to found all sources that we can use to gather “English – sign writing” combination sentences for our dataset.

In first try we gathered 1500 words and small sentences combinations, and we use it to build our first experimental model.

We use a sequence to sequence (seq2seq) model and after training the model we were able to input an English word, such as \*"¿translator ", and return the sign:

\*" 𝠀񀀁񀀩񆙡񋎩񋎽񂈁񂈉񆿅񆿕񋸥𝠃𝤨𝥇񆙡𝣪𝣟񀀁𝤅𝣕񀀩𝣿𝣚񋎽𝣥𝣱񋎩𝣦𝣆񂈁𝤚𝤌񂈉𝤁𝤌񆿅𝤛𝤮񆿕𝤀𝤮񋸥𝤍𝥁 "

Which is the representation of word “translator” in American Sign Writing.

# To train our model we used Tensor flow in “google Colab” and we followed steps in “Neural Machine Translation with Attention” Notebook available at:

# <https://colab.research.google.com/github/tensorflow/docs/blob/master/site/en/r2/tutorials/text/nmt_with_attention.ipynb>

The steps are as follow:

**A: prepare the dataset:**

We stored the dataset in a “.txt” file in Unicode format, every line in this file is an “English + American sign language” sentence and they are separated with tab “/t”

Example:

𝠀񀀁񀀉񈗥񈗵񋸦𝠃𝤝𝤨񀀁𝤎𝣤񀀉𝣰𝣮񈗥𝤏𝤇񈗵𝣱𝤑񋸦𝤆𝤜 ¿come

1. Add a \*start\* and \*end\* token to each sentence.

2. Clean the sentences by removing special characters.

3. Create a word index and reverse word index (dictionaries mapping from word → id and id → word).

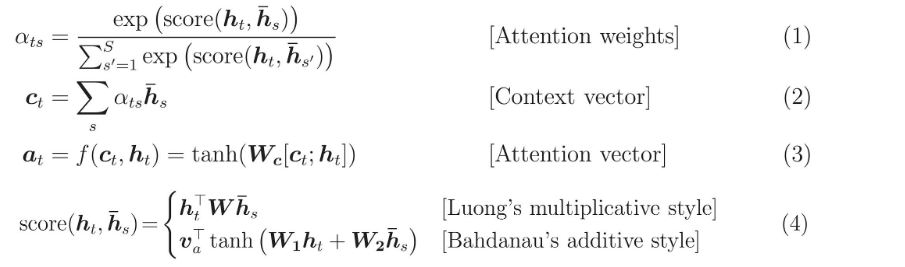
4. Pad each sentence to a maximum length.

### **B: Create a tf.data dataset**

## **C: Write the encoder and decoder model:**

The input is put through an encoder model which gives us the encoder output of shape (batch\_size, max\_length, hidden\_size) and the encoder hidden state of shape (batch\_size, hidden\_size).

Here are the equations that are implemented:



**Figure 4.1 Encoder and Decoder formula**

If we consider this notation:

* FC = Fully connected (dense) layer
* EO = Encoder output
* H = hidden state
* X = input to the decoder

## **D: Define the optimizer and the loss function:**

## **E: Checkpoints (Object-based saving)**

## **F: Training:**

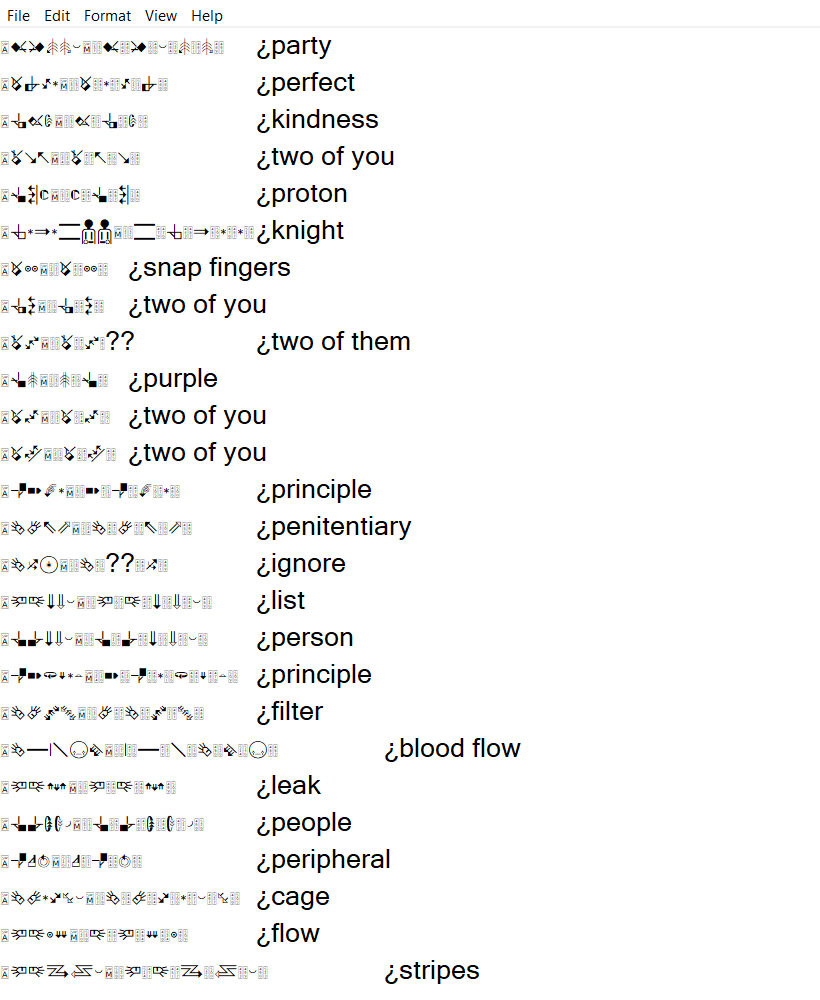
1. Pass the input through the encoder which return encoder output and the encoder hidden state.
2. The encoder output, encoder hidden state and the decoder input (which is the start token) is passed to the decoder.
3. The decoder returns the predictions and the decoder hidden state.
4. The decoder hidden state is then passed back into the model and the predictions are used to calculate the loss.
5. We Used teacher forcing to decide the next input to the decoder.
6. Teacher forcing is the technique where the target word is passed as the next input to the decoder.
7. The final step is to calculate the gradients and apply it to the optimizer and back propagate.

## **G: Translate:**

* The evaluate function is similar to the training loop, except we don't use teacher forcing here. The input to the decoder at each time step is its previous predictions along with the hidden state and the encoder output.
* Stop predicting when the model predicts the end token.
* And store the attention weights for every time step.

Note: The encoder output is calculated only once for one input.

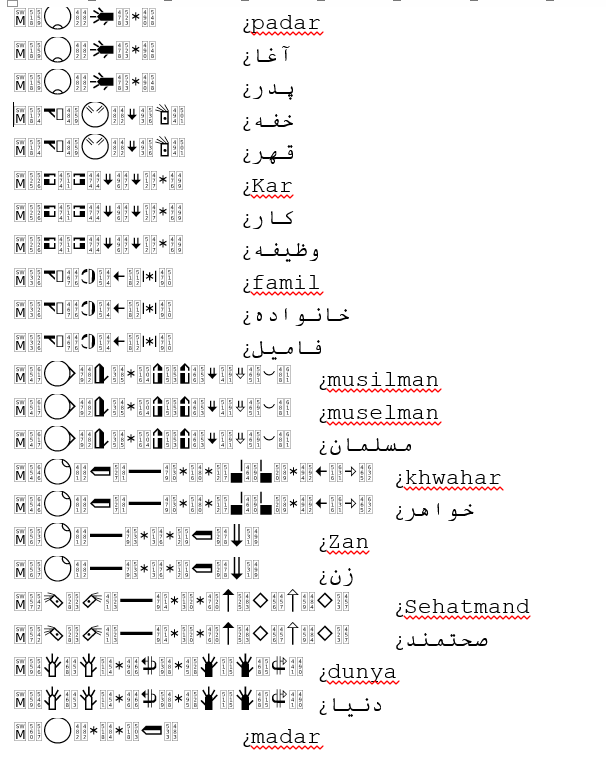
**Here is the sample of prepared file format for training our data:**

****

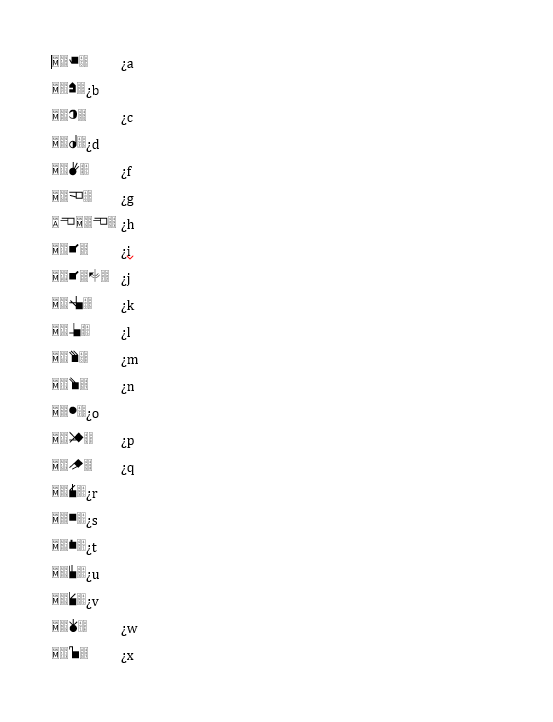
**Figure 4.2 Dataset sample**

****

**Figure 4.3 Dataset sample**

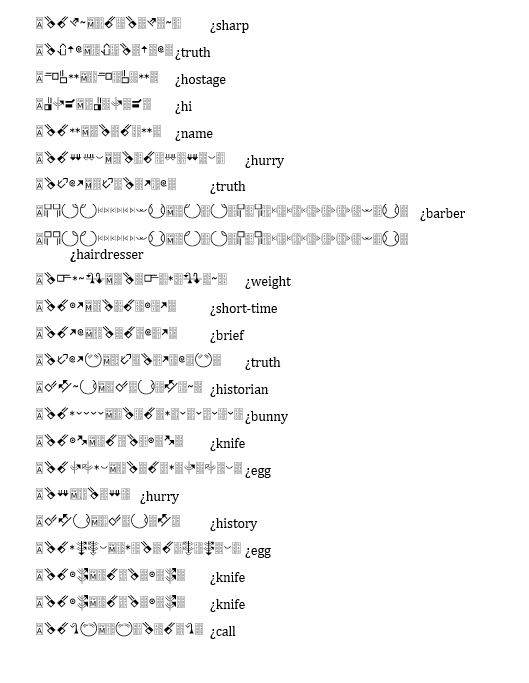
****

**Figure 4.4 Dataset sample**

****

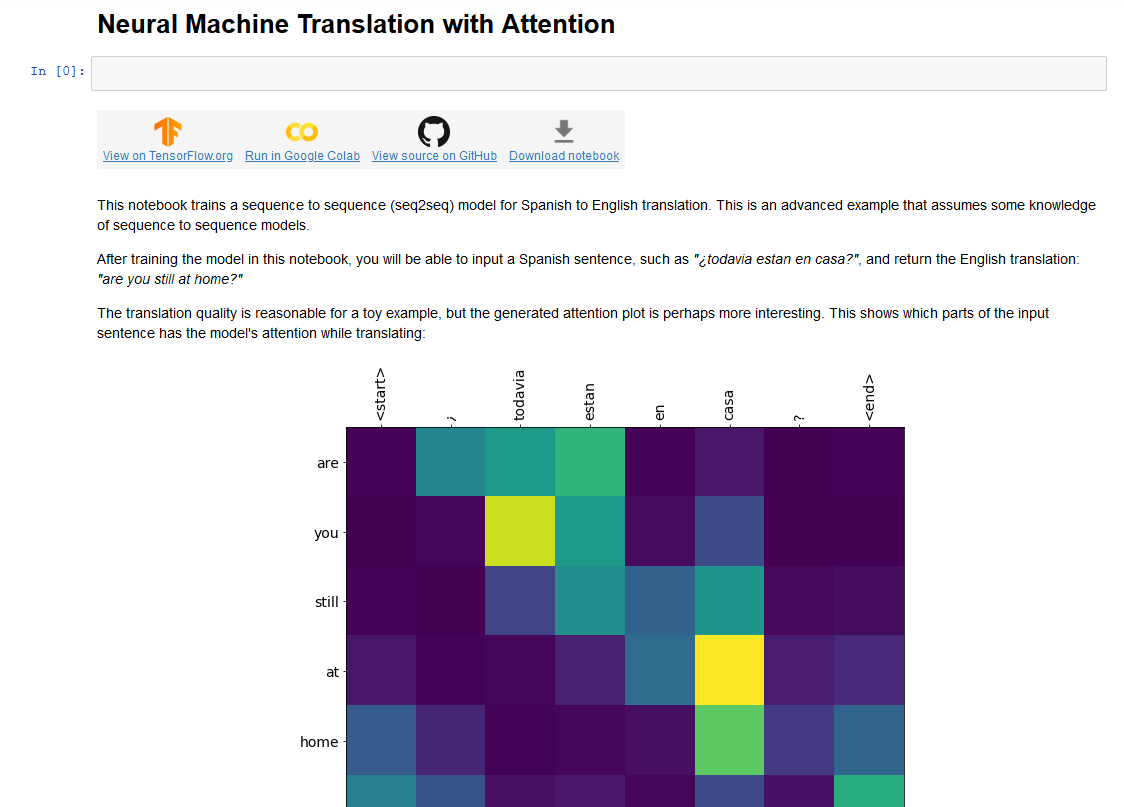
**Figure 4.5 Dataset sample**

** Figure 4.6 Dataset sample**

****

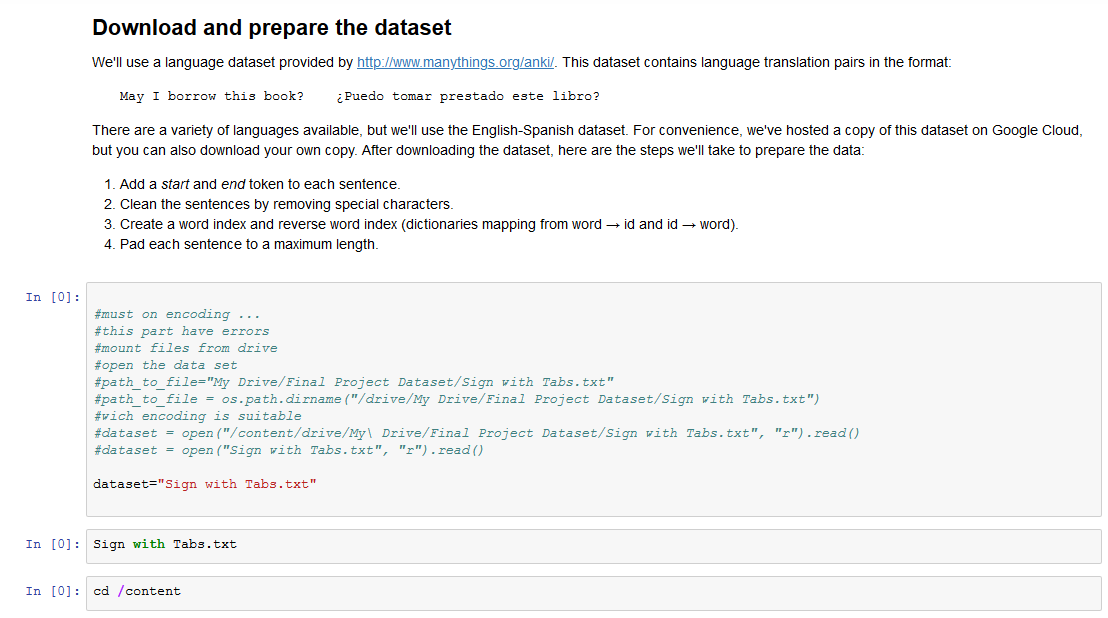
**Figure 4.7 Dataset sample**

**Here are screenshots of our trained/tested model:**

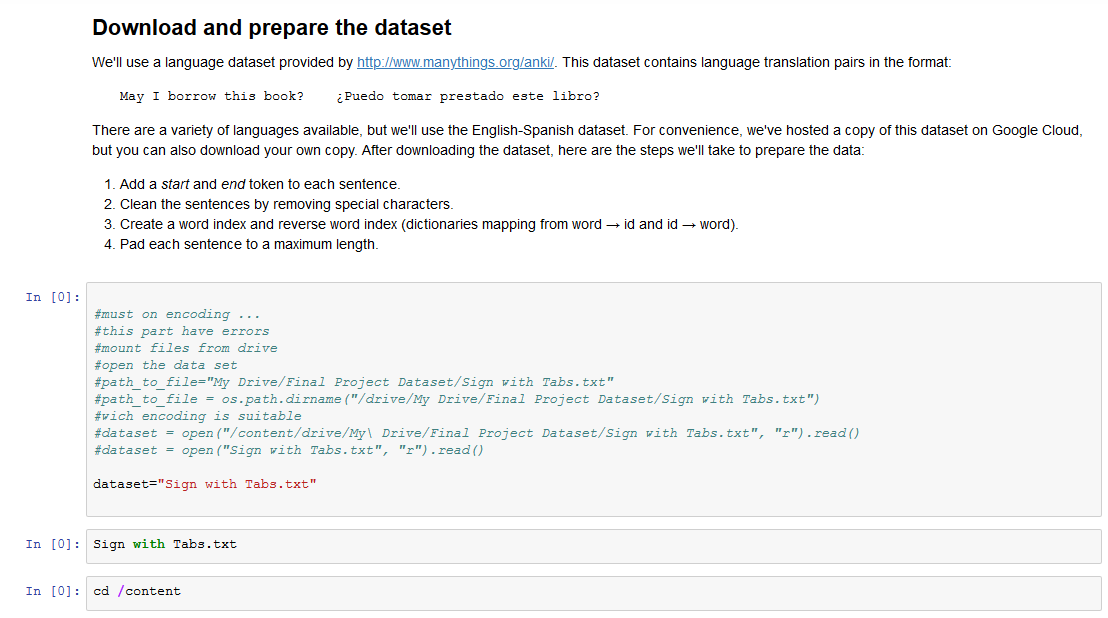


**Figure 4.8 Preparing the Model**





**Figure 4.9 preparing the model for train using TenserFlow**



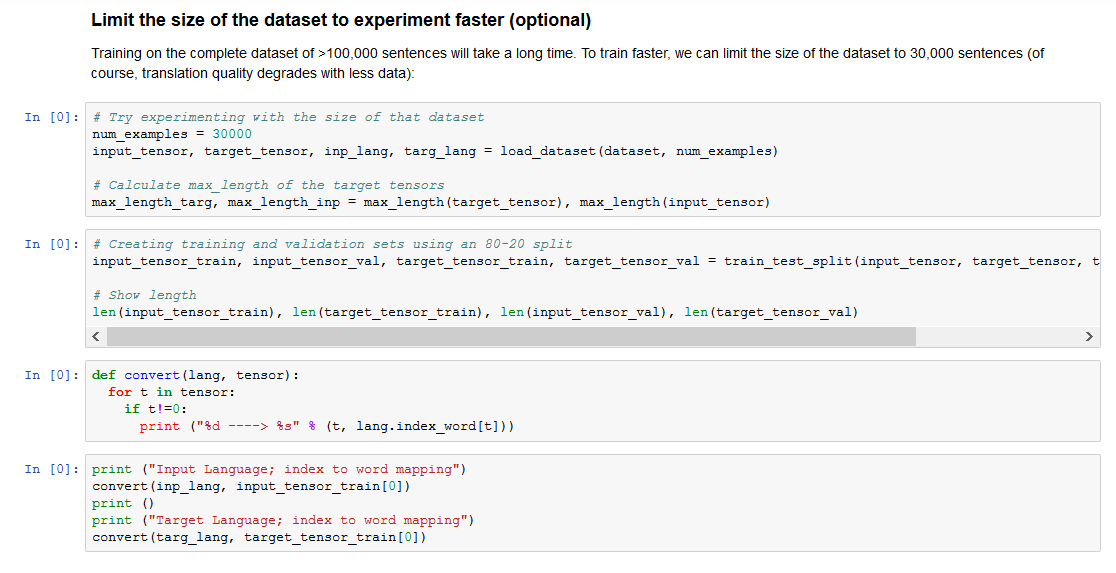
**Figure 4.10 Preparing Dataset for train**



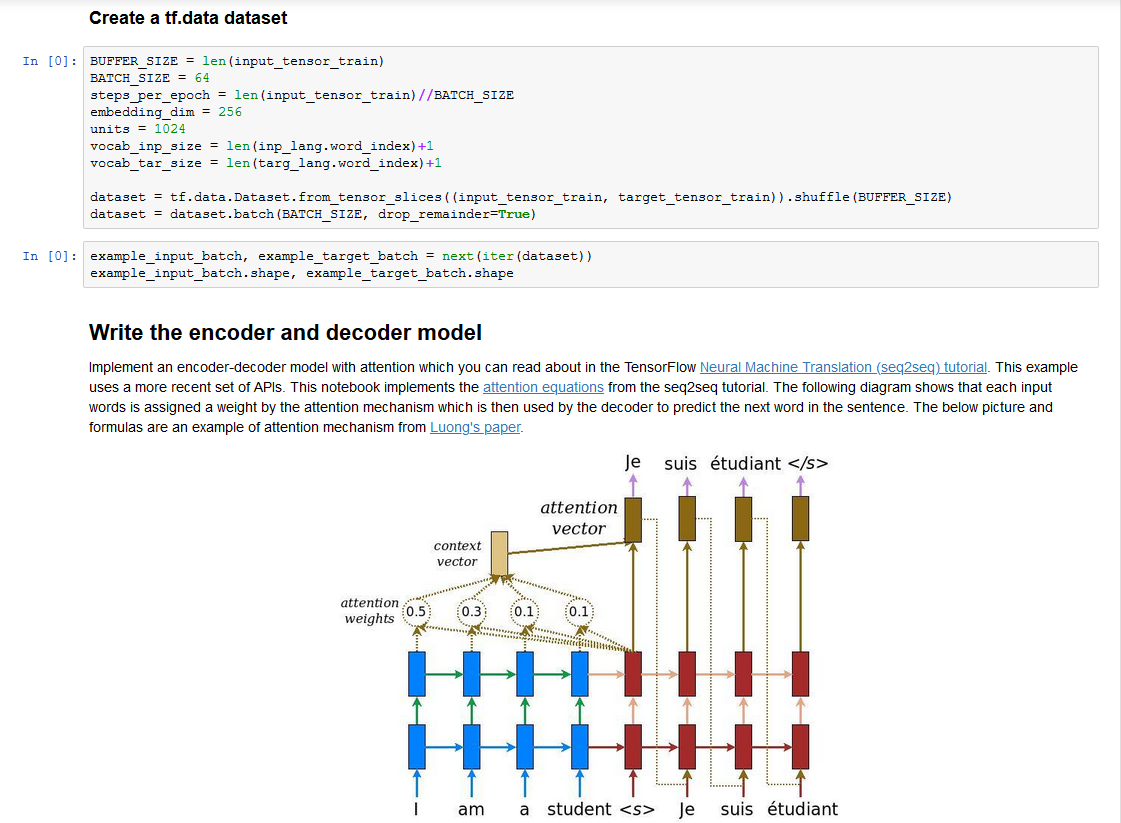
**Figure 4.11 Reading files from directory**



**Figure 4.12 Perform normalization on dataset**



**Figure 4.13 Perform normalization on dataset**



**Figure 4.14 RNN encoder and decoder for Se2Seq model**



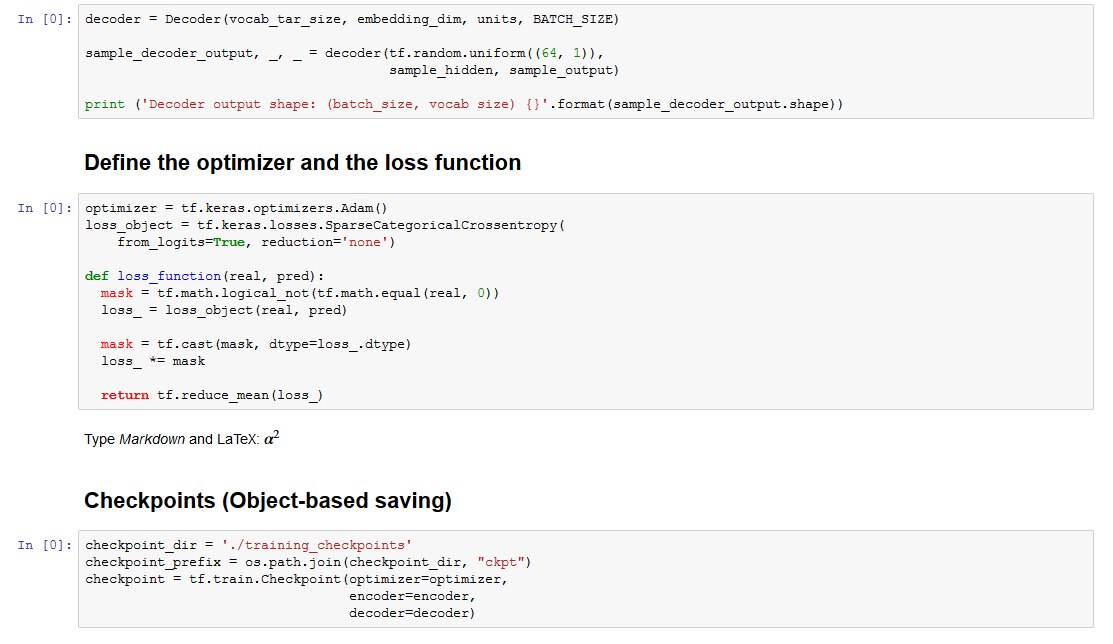
**Figure 4.15 Encoder function**



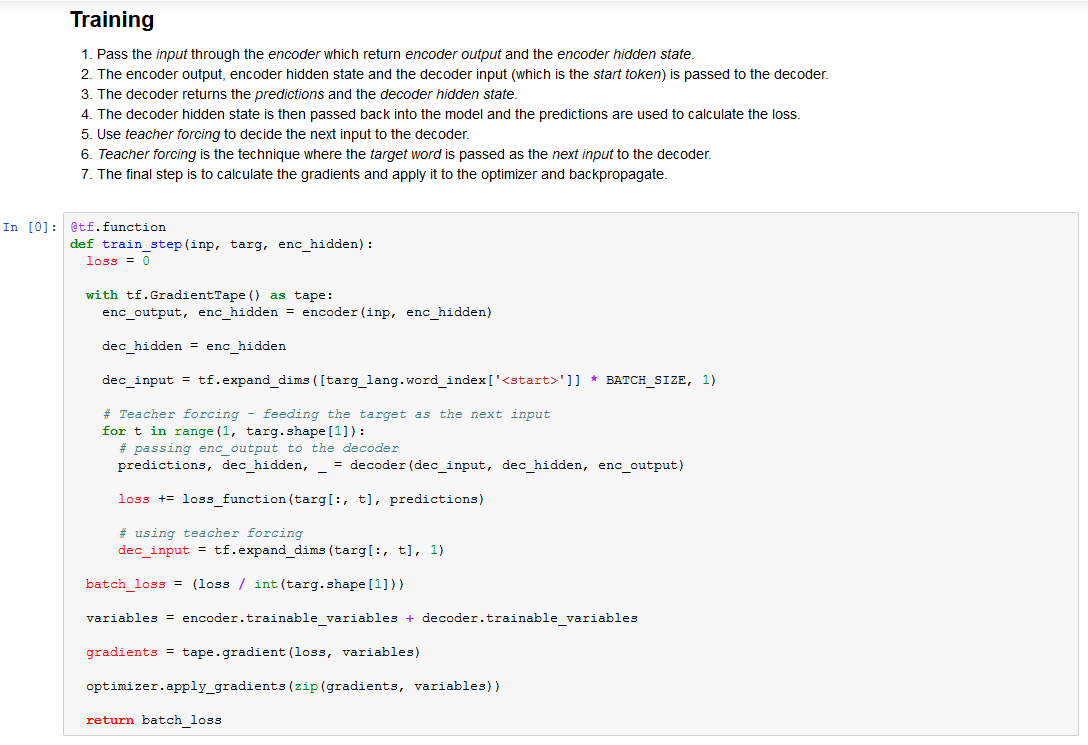
**Figure 4.16 Encoder function**



**Figure 4.17 Decoder function**



**Figure 4.18 Define Optimizer and loss function**



**Figure 4.19 Train the model**



**Figure 4.20 calculating the total loss**



**Figure 4.21 evaluating the model**



**Figure 4.22 Test the model**

**Chapter 4**

**4.1 Result and Discussion**

We proved technology is a way that ordinary people is able to communicate with Deaf and remove that gap between them by using an automatic system that can translate English words and sentences into international sign language just like other machine translations e.g.: Google translate. During literature review we realized what makes a system better for deaf in order for better communication with them, there are a lot of machine learning methods exists for converting one language to sign language such us RNN algorithm Seq2Seq model, TenserFlow library for machine translation and some gaming engines such us Unity that can be used in making an avatar for sign movement.

As we discussed our three Research questions we have found answers according to our literature reviews:

1. How our solution will reduce the gap between Deaf and people?

By making a platform that can exchange communication between a deaf and ordinary person. We studied some of existing systems such as: Tessa this platform was designed for communication between a deaf and post-office employee, or another system like HandTalk which was translating Spanish language into Sign Language, but we claimed that we have another approach which removes the gap between a deaf and ordinary person that was using of International Sign writing this approach was good among all as it’s scalable and be integrated with any other language.

1. Which ways are convenient to include deaf in technology?

According to our research studies, we have found that there are many ways to include deaf in technology like: smart glove [9], sound transaction tool [16] and so on, but we searched and found a proper way that is Digital and Automation translation. This mechanism is nothing more but translating of one language to Sign language and displaying the translated sequences of words using of animation tools like: SGML, Unity or another.

1. Which Models can be used to automate a system for communication between ordinary people and Deaf?

As we studied many systems have used different models for making communication automated like Scribe4Me, it used Touch Screen API for translation of languages. Moreover, Deaf-Mute Communication Interpreter this system used sensor and IOT devices for capturing each movement of a deaf, but we have used Deep Learning models which are the newest. We have used Tensorflow a machine learning library which is used for machine translation, deeper we used RNN algorithm Seq2Seq model for training our data and predicting it with minimum loss function.

**Chapter 5**

**5.1 Conclusion**

Deaf are people who can’t talk and hear, and they can’t recognize our thoughts, thus Deaf think there is a gap between them and ordinary people. Therefore Ava is an experimental system that aims to aid communication between a deaf person and an ordinary person by translating the ordinary person’s speech to sign language.

We conclude that the dataset is not a complete space, and there are still other feature vectors missing from it.Our goal in developing this trial system was to establish whether the introduction of a limited speech-to-sign translation system for ordinary people would be beneficial to deaf whose primary means of communication was sign language. Moreover, we used International sign writing mechanism but for testing our first version of our system we are using American Sign Language, in future it is possible to fit any sign wiring language with our system. In addition, concatenation of signing is more fluent and controlled for avatar than for video signing, as the exact positioning of the avatar can be manipulated. For these reasons, we decided to display the signs using an avatar (3D character).

**5.2 Future Works**

While our claim is almost done, there is still room for making the project fully localized for deaf of Afghanistan also more interactive and friendly, as well as developing more signs for deaf for better accuracy of system and its interaction.

There is still room for adding a functionality that deaf also would be able to deliver his/her understanding to ordinary people.

**Acknowledgments**

Our thank goes to Allah the almighty who gave us the ability, knowledge and all the skills to accomplish this project. We are grateful to Allah for his support during the tough and stressed moments of the work. We would like thank our supervisor Asst.Prof Rafiullah Momand, for his academic support and valuable guidance throughout this project, He helped us a lot and we really enjoyed working under his supervision. And we highly appreciate computer science faculty for giving us this opportunity to approve our claim.

**Bibliography**

*[1] Kato, Mihoko. "A study of notation and sign writing systems for the deaf." Intercultural Communication Studies 17, no. 4 (2008): 97-114.*

*[2] B. Mazor and B. L. Zeigler. The design of speech-interactive dialogs for transaction automation systems. Speech Communication,*

*17:313–320, November 1995.*

*[3] R. Conrad. The deaf school child. Harper and Row, 1979.*

*[4] Michael Saar and Helena Arthur‐Okor (2013) "Reference services for the deaf and hard of hearing"*

*[5] Maiorana-Basas, Michella, and Claudia M. Pagliaro. "Technology use among adults who are deaf and hard of hearing: A national survey." Journal of deaf studies and deaf education 19.3 (2014): 400-410.*

*[6] Al-Fityani, Kinda, and Carol Padden. "Sign language geography in the Arab world." Sign languages: A Cambridge survey (2010): 433-450.*

*[7] D. Wood, H. Wood, A. Griffiths, and I. Howarth. Teaching and talking with deaf children. John Wiley and Sons, 1986.*

*[8] Cox, S., Lincoln, M., Tryggvason, J., Nakisa, M., Wells, M., Tutt, M. and Abbott, S., 2002, July. Tessa, a system to aid communication with deaf people. In Proceedings of the fifth international ACM conference on Assistive technologies (pp. 205-212). ACM.*

*[9] U. Gollner et al., "Mobile Lorm glove introducing a communication device for deaf-blind people",* Proc. 6th Int. Conf. Tangible Embedded and Embodied Interaction*, pp. 127-130, 2012.*

*[10] Rajamohan, Anbarasi, R. Hemavathy, and M. Dhanalakshmi. "Deaf-mute communication interpreter." International Journal of Scientific Engineering and Technology 2.5 (2013): 336-341.*

*[11]* Khaled Assaleh, M. Al-Rousan. Recognition of Arabic Sign Language Alphabet Using Polynomial Classifiers. EURASIP Journal on Applied Signal Processing 2005:13, 2136–2145

*[12] Matt Huenerfauth. Generating American Sign Language Classifier Predicates For English-To Asl Machine Translation. Ph.D dissertation, University of Pennsylvania, Department of Computer and Information Science, Philadelphia, PA, USA 2006*

*[13] Halawani, Sami M. "Arabic sign language translation system on mobile devices." IJCSNS International Journal of Computer Science and Network Security 8.1 (2008): 251-256.*

*[14] Rodríguez, M.A. (1991). “Lenguaje de signos” Phd Dissertation. Confederación Nacional de Sordos Españoles (CNSE) and ONCE. Madrid. Spain. 1991.*

*[15] San-Segundo, R., Barra, R., D'Haro, L.F., Montero, J.M., Córdoba, R. and Ferreiros, J., 2006. A spanish speech to sign language translation system for assisting deaf-mute people. In Ninth International Conference on Spoken Language Processing.*

*[16] Matthews, T., Carter, S., Pai, C., Fong, J. and Mankoff, J., 2006, September. Scribe4Me: Evaluating a mobile sound transcription tool for the deaf. In International Conference on Ubiquitous Computing (pp. 159-176). Springer, Berlin, Heidelberg.*

*[17*] Selwyn, N., 2015. Data entry: towards the critical study of digital data and education. Learning, Media and Technology, 40(1), pp.64-82.

*[18] S. Morrissey, A. Way, D. Stein, J. Bungeroth, andH. Ney(2007), “Towards a hybrid data-driven mt systemfor sign languages,” in Machine Translation Summit, Copenhagen, Denmark, Sept. 2007, pp. 329–335.*

*[19] Barberis, D., Garazzino, N., Prinetto, P., Tiotto, G., Savino, A., Shoaib, U. and Ahmad, N., 2011, November. Language resources for computer assisted translation from italian to italian sign language of deaf people. In Proceedings of Accessibility Reaching Everywhere AEGIS Workshop and International Conference, Brussels, Belgium (November 2011).*

*[20] Michael Auli, Michel Galley, Chris Quirk, and Geoffrey Zweig. 2013. Joint language and translation modeling with recurrent neural networks. In Proceedings of the ACL Conference on Empirical Methods in Natural Language Processing (EMNLP), pages 1044–1054.*

*[21] Amittai Axelrod, Xiaodong He, and Jianfeng Gao. 2011. Domain adaptation via pseudo in-domain data selection. In Proceedings of the ACL Conference on Empirical Methods in Natural Language Processing (EMNLP), pages 355–362.*

*[22] Fr´ed´eric Bastien, Pascal Lamblin, Razvan Pascanu, James Bergstra, Ian J. Goodfellow, rnaud Bergeron, Nicolas Bouchard, and Yoshua Bengio. 2012. Theano: new features and speed improvements. Deep Learning and Unsupervised Feature*

*Learning NIPS 2012 Workshop.*

*[23]Sriram, A., Jun, H., Satheesh, S. and Coates, A., 2017. Cold fusion: Training seq2seq models together with language models. arXiv preprint arXiv:1708.06426.*

*[24]Vaswani, A., Bengio, S., Brevdo, E., Chollet, F., Gomez, A.N., Gouws, S., Jones, L., Kaiser, Ł., Kalchbrenner, N., Parmar, N. and Sepassi, R., 2018. Tensor2tensor for neural machine translation. arXiv preprint arXiv:1803.07416.*

*[25]Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard,M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D. G., Steiner, B., Tucker, P., Vasudevan,V., Warden, P., Wicke, M., Yu, Y., and Zheng, X. (2016). Tensorflow: A system for large-scale machinelearning. In12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)*

[*26] Axon, Samuel (September 27, 2016).*[*"Unity at 10: For better—or worse—game development has never been easier"*](https://arstechnica.com/gaming/2016/09/unity-at-10-for-better-or-worse-game-development-has-never-been-easier/)*. Ars Technica.*[*Archived*](https://web.archive.org/web/20181005025906/https:/arstechnica.com/gaming/2016/09/unity-at-10-for-better-or-worse-game-development-has-never-been-easier/)*from the original on October 5, 2018. Retrieved October 17, 2018[27]Riccitiello, John (October 23, 2014).*[*"John Riccitiello sets out to identify the engine of growth for Unity Technologies (interview)"*](https://venturebeat.com/2014/10/23/john-riccitiello-sets-out-to-identify-the-engine-of-growth-for-unity-technologies-interview/)*. VentureBeat (Interview). Interviewed by Dean Takahashi.*[*Archived*](https://web.archive.org/web/20150117174953/http:/venturebeat.com/2014/10/23/john-riccitiello-sets-out-to-identify-the-engine-of-growth-for-unity-technologies-interview/)*from the original on January 17, 2015. Retrieved*

*[28]* Kanter, J.M. and Veeramachaneni, K., 2015, October. Deep feature synthesis: Towards automating data science endeavors. In *2015 IEEE International Conference on Data Science and Advanced Analytics (DSAA)* (pp. 1-10). IEEE.

**Appendix A**

**Glossary**

**A**

**Avatar:** an avatar is the graphical representation of the user or the user's alter ego or character.

**ASL (American Sign Language):** American Sign Language is a natural language that serves as the predominant sign language of Deaf communities in the United States and most of Anglophone Canada

**ASR (Automatic Speech Recognition):** Automatic speech recognition (ASR) is the process and the related technology for converting the speech signal into its corresponding sequence of words or other linguistic entities by means of algorithms implemented in a device, a computer, or computer clusters**.**

**AM (Acoustic Model):** is used in [automatic speech recognition](https://en.wikipedia.org/wiki/Automatic_speech_recognition) to represent the relationship between an [audio signal](https://en.wikipedia.org/wiki/Audio_signal) and the [phonemes](https://en.wikipedia.org/wiki/Phonemes) or other linguistic units that make up speech. The model is learned from a set of audio recordings and their corresponding transcripts. It is created by taking audio recordings of speech, and their text transcriptions, and using software to create statistical representations of the sounds that make up each word.

**B**

**BSL (British Sign Language):** British Sign Language is a sign language used in the United Kingdom, and is the first or preferred language of some deaf people.

**C**

**CPU:** British Sign Language is a sign language used in the United Kingdom, and is the first or preferred language of some deaf people.

**D**

**Deaf:** Deaf culture is the set of social beliefs, behaviors, art, literary traditions, history, values, and shared institutions of communities that are influenced by [deafness](https://en.wikipedia.org/wiki/Deafness) and which use [sign languages](https://en.wikipedia.org/wiki/Sign_language) as the main means of communication.

**deaf:** A deaf person has little to no hearing. Hearing loss may occur in one or both ears. In children, hearing problems can affect the ability to learn spoken [language](https://en.wikipedia.org/wiki/Language) and in adults it can create difficulties with social interaction and at work.

**F**

**FSW:** Formal SignWriting is a faithful character-encoding of Sutton SignWriting based on 2-dimensional mathematics. FSW defines a formal language for written sign languages where any sign of any sign language can be written as a string of ASCII or Unicode characters.

**G**

**GPU:** Formal SignWriting is a faithful character-encoding of Sutton SignWriting based on 2-dimensional mathematics. FSW defines a formal language for written sign languages where any sign of any sign language can be written as a string of ASCII or Unicode characters.

**H**

**HCI:** Human-computer interaction is a multidisciplinary field of study focusing on the design of computer technology and, in particular, the interaction between humans (the users) and computers.

**I**

**IMWA:** International Movement Writing Alphabet used to record all human and animal gestures.

**J**

**JSL:** Japanese Sign Language, also known by the acronym JSL, is the dominant sign language in Japan and is a natural language, like the spoken Japanese language for deaf.

**L**

**LIS:** Italian Sign Language or LIS is the visual language used by deaf people in Italy.

**LSTM:** Long short-term memory is an artificial recurrent neural network architecture used in the field of deep learning. Unlike standard feed forward neural networks, LSTM has feedback connections. It can not only process single data points, but also entire sequences of data.

**M**

**MOSES:** Moses is a free software, statistical machine translation engine that can be used to train statistical models of text translation from a source language to a target language. Moses then allows new source-language text to be decoded using these models to produce automatic translations in the target language.

**N**

**NMT (Neural Machine Translation):** is a subfield of computational linguistics that is focused on translating text from one language to another**.**

**P**

**PC:** Personal computer

**PDA:** A personal digital assistant, also known as a handheld PC, is a variety mobile device which functions as a personal information manager.

**R**

**RNN:** A recurrent neural network is a class of artificial neural networks where connections between nodes form a directed graph along a temporal sequence. Unlike feed forward neural networks, RNNs can use their internal state (memory) to process sequences of inputs.

**S**

**SMS:** (short message service) is a text messaging service component of most telephone, Internet, and mobile device systems. It uses standardized communication protocols to enable mobile devices to exchange short text messages.

**SGML: (**Standard generalized markup Language) is a standard for how to specify a document markup language or tag set. Such a specification is itself a document type definition (DTD). SGML is not in itself a document language, but a description of how to specify one. It is metadata.

**SSL:** Spanish Sign Language is a sign language used mainly by deaf people in Spain and the people who live with them.

**Seq2Seq:** A Seq2Seq model is a model that takes a sequence of items (words, letters, time series, etc) and outputs another sequence of items.

**T**

**TTS:** Text to speech is a form of speech synthesis that converts text into spoken voice output. Text to speech systems were first developed to aid the visually impaired by offering a computer-generated spoken voice that would "read" text to the user.

**T2T (Tensor 2 Tensor):** is a library of deep learning models and datasets designed to make deep learning more accessible and [accelerate ML research](https://research.googleblog.com/2017/06/accelerating-deep-learning-research.html).

**Appendix B**

**Interviews**

Requirement Gathering for Ava system was in Kabul, Afghanistan and USA deaf. We had interviews with Ministry of Education and Deaf School, also with deaf students and we consider how they are communicating with each other. It helped us to find and gather those information which were need for making this system.

**B.0.1 People Met for discussion**

1. Mr. Mohammad Hasib Amini
2. Mr. Irfan Safi
3. Ms.Fatana Dawoodi

**B.1 My interview with Mr.Mohammad Hasib Aimini**

Q: what is your name and your job?

A: my name is Mohammad Hasib amini and I am Project manager specialist at Ministry of Education.

Q: Do you face any problem when a deaf is coming to your office?

A: sure, I don’t know how to tell him/her what I need, he/she doesn’t understand my thoughts.

Q: what do you want?

A: well, I want some easy way that I could be able to communicate with a Deaf.

**B.1 My interview with Mr.Irfan Safi**

Q: what is your name and your job?

A: my name is Irfan Safi and I am Statistics specialist at Ministry of Education.

Q: Ok, how many deaf we have in Afghanistan?

A: around 30% of Afghanistan people are deaf.

Q: how many deaf schools and students we have in Afghanistan?

A: well, we have more than 20 deaf schools in Afghanistan.

**B.1 My interview with Ms.Fatana Dawoodi**

Q: what is your name and your job?

A: my name is Fatana Dawoodi and I am manager at Kabul deaf high school.

Q: how many students you have in school?

A: we have 400 deaf students half of them are in the morning shift and half of them are afternoon.

Q: how your teachers teaching them?

A: our teachers are also deaf they are communicating with them through signs.

Q: can you deliver your speech to them?

A: actually no, I am getting help from a translator.

Q: do you want a system that can help you instead of a translator?

A: sure, we defintly need that kind.

**Declaration**

This thesis is a presentation of my original research work. Wherever contributions of others are involved, every effort is made to indicate this clearly, with due reference to the literature, and acknowledgment of collaborative research and discussions. The work was done under the guidance of Asst.Prof Rafiullah Momand, at the Department of Information Systems, Computer Science Faculty, Kabul University, and Kabul.

Ahmad Zia Yosufi........................................................

Mehreen Najm...............................................................

Name(s) and Signature(s)

In my capacity as supervisor of the candidate’s thesis, I certify that the above

Statements are true to the best of my knowledge.

.....................................................................................................................

Rafiullah Momand, December 2019