**Name: Mohamed Ahmed Mohamed Ahmed El Harakani / ID:20201442457**

**Name:Youmna Ayman Abdo El Komy / ID: 2003120**

**Name:Georgino Nabil / ID: 2003121**

**NLP Project**

Our project is a machine translation project we have three model two models translate English to French and the third model translate hieroglyphic to English

First of all we are going to talk about data and data preprocessing

Our dataset is English to French dataset and the shape of data I (229803,)

We are going now to preprocess our text by removing punctuation and to lower case our data

Then we have a function call preprocess\_target\_data this function created to identify to the French data were is the start and the end of the sentence

Then splitting data set we did a function that takes data\_source and data\_target and size that we want to split then split our data int train and validation

After splitting we use tokenization to tokenize our data we use two tokenizer one for data\_source and the other for data\_target

Then to the First model:

A diagram of a data flow

Description automatically generated

We have encoder and decoder the encoder take English data then return the final values of hidden state and cell state these values we take it to be the initial value for the lstm for the decoder which take as input the French data then pass the output from lstm of decoder to a dense layer then get the output sentence

Model evaluation

A screenshot of a computer

Description automatically generated

Model Blue score

A screenshot of a computer code

Description automatically generated

Plot of Val\_Accuracy and training\_Accuracy

A graph with blue line and orange line

Description automatically generated

Plot of Training\_loss and Val\_Loss

A graph of loss and validation

Description automatically generated

**Model 2 The Attention Model:**

In the attention model we have the encoder and the decoder

The difference between this model and the previous model is the attention model weighted the encoder output that mean that now we know which input word have the most wait to predict the word in French

That is the main difference between the attention model and the previous model

The loss of this model is 0.15

We can say other thing that attention model is better to use in long sequence sentence that using the first model

**The hierographic translation model:**

In this model we are using the cnn and nlp we first train cnn model return for us the code of each small picture of a character

And then we path this code to a transformer that is trained on the same concept of code then we can get the result of the transelation