**BBM467 - Blog Post Project**

*Jaguar*

*21946003, Oğuzhan Çetinkaya*

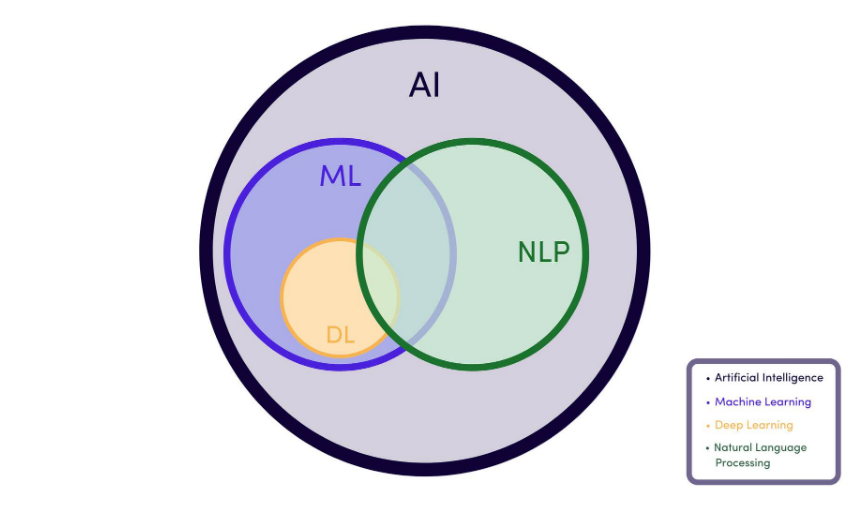
*21727355,Meryem Ülkü Kara*

**How does Google Translate work with Natural Language Processing(NLP)?**

Google translate is a successful translation application that we all need for certain reasons and is developing day by day. During its work, it provides various services such as audio, written, document, and image. If we think about Google Translate and its services, we realize that every language is very different from the others and all of them have various regional differences even within themselves, there are differences in tone of voice and speaking speed between people. Even in a sentence, a word can have more than one meaning, and what we want to say can change according to our intonation. Despite all these differences and difficulties, Google Translate offers us a successful service that is constantly improving.

Google Translate is a machine, it cannot understand and interpret like humans, but how does it offer these services to us? Google Translate uses Natural Language Processing to overcome differences in language features and achieve successful results. I will explain what Natural Language Processing is and how it is used in Google translate.

**Natural Language Processing (NLP)**

Natural language processing (NLP) refers to the branch of computer science—and more specifically, the branch of artificial intelligence or AI—concerned with giving computers the ability to understand text and spoken words in much the same way human beings can. Natural Language Processing is the technology used to aid computers to understand natural human language. This commonly includes detecting sentiment, machine translation, or spell check - often repetitive but cognitive tasks. Through NLP, computers can accurately apply linguistic definitions to speech or text.

NLP is a branch of Artificial Intelligence & Machine Learning

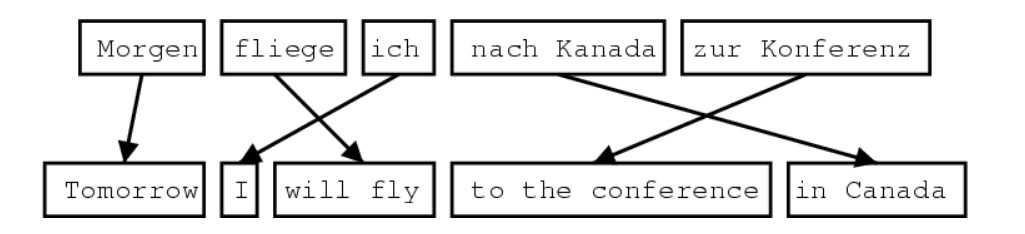
Several NLP tasks break down human text and voice data in ways that help the computer make sense of what it's ingesting. Some of these tasks include the following:

* Spell Checking: It is the language and vocabulary check of a written sentence. (e.g. Microsoft word, Grammarly)
* Sentiment Analysis: It is the determination of mood and emotions in a particular text.
* Part Of Speech Tagging (Grammatical Tagging): It is the process of determining the part of speech of a particular word or piece of text based on its use and context. Part of speech identifies ‘make’ as a verb in ‘I can make a paper plane,’ and as a noun in ‘What make of car do you own?’
* Word-sense disambiguation (WSD): It is the process of identifying which sense of a word is meant in a sentence or other segment of context.
* Named-entity recognition (NER, Entity Identification): It is a subtask of information extraction that seeks to locate and classify named entities mentioned in unstructured text into pre-defined categories.
* Co-reference resolution: It is the task of identifying if and when two words refer to the same entity.
* Natural language generation(NLG): It is a software process that produces natural language output.

There are four different types of machine translation in NLP:

* Statistical machine translation (SMT): It is a machine translation approach that uses large volumes of bilingual data to find the most probable translation for a given input. Statistical machine translation systems learn to translate by analyzing the statistical relationships between original texts and their existing human translations. The most important components in statistical machine translation are the translation model and the language model.
* Rules-Based Machine Translation (RBMT): RBMT systems were the first commercial machine translation systems and are based on linguistic rules that allow the words to be put in different places and to have different meanings depending on the context. Rules are developed by human language experts and programmers who have deployed extensive efforts to understand and map the rules between two languages.
* Hybrid Machine Translation (HMT): HMT, as the term demonstrates, is a mix of RBMT and SMT. The motivation for developing hybrid machine translation systems stems from the failure of any single technique to achieve a satisfactory level of accuracy.
* Neural Machine Translation (NMT, Deep Neural Machine Translation): It is a state-of-the-art machine translation approach that utilizes neural network techniques to predict the likelihood of a set of words.

**Google Translate with NLP**

Google Translate was developed in April 2006 using Statistical Machine Translation, a type of NLP. The working principle of STM is as follows:

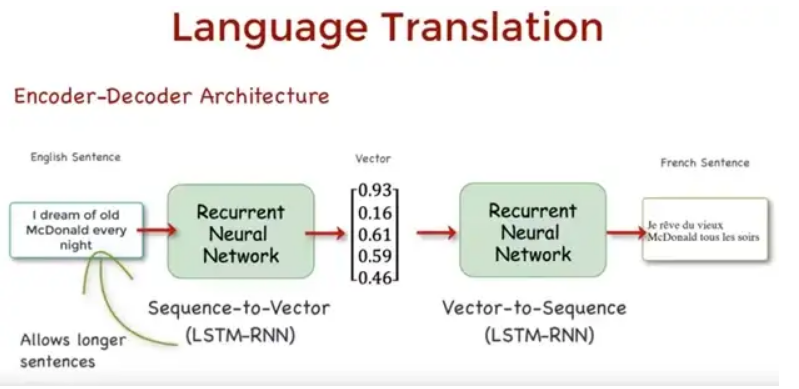
1. Separating the sentence in the source language into phrases
2. Searching for each of these phrases in the Phrase Table/Translation Model and generating their target language translations
3. Engineer then uses the Phrase Table/Translation Model to reorder these patterns to optimize the Fluency of the translation.

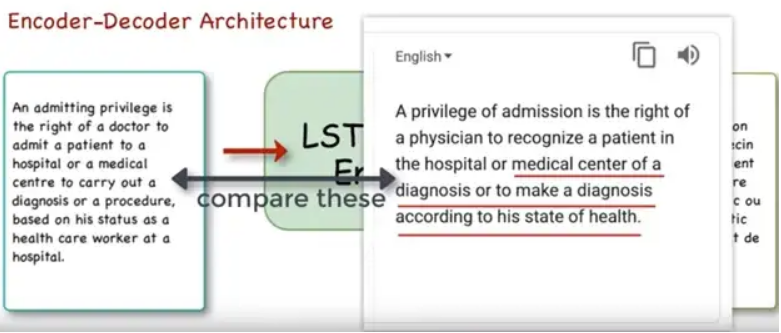
In this principle, there were many deficiencies such as the thought that the sentence wanted to express, choosing from many meanings of the word used.

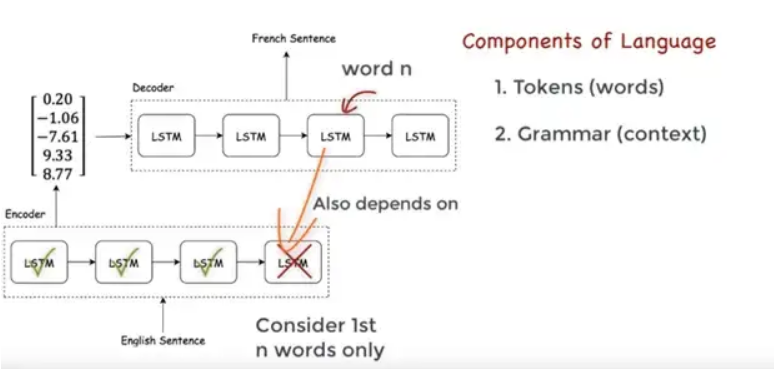
In September 2016, the Google Neural Machine Translation system (GNMT) system, based on a powerful neural network with deep learning capabilities, replaced Statistical Machine Translation and replaced SMT. Neural machine translation uses neural networks to translate source text to target text, and neural networks can work with very large datasets and require little supervision. Neural machine translation systems have two main sections: an encoder network and a decoder network. Both are neural networks. A neural network is an interconnected series of nodes, loosely modeled on the human brain. It’s an information system in which input data is passed through these nodes to produce output. This neural network architecture is called a sequence-to-sequence neural network (Seq2Seq), which works by looking at one source-language sentence and producing a corresponding target-language sentence.

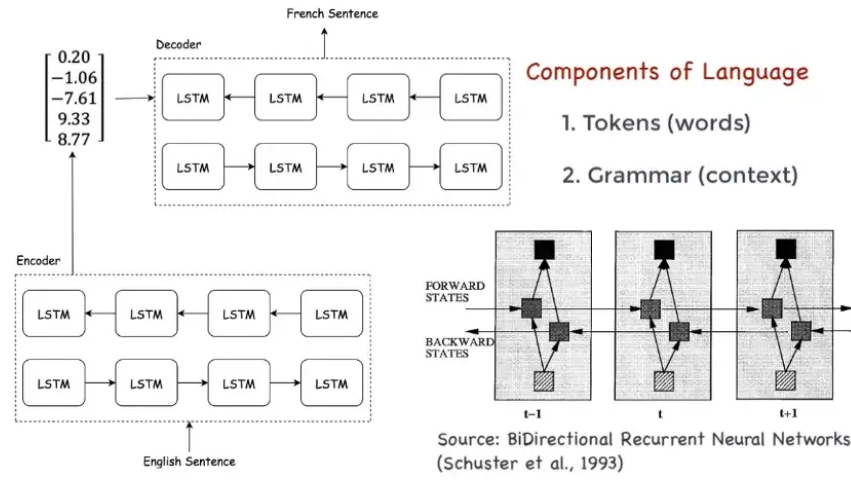
I'll try to cover how it all works:

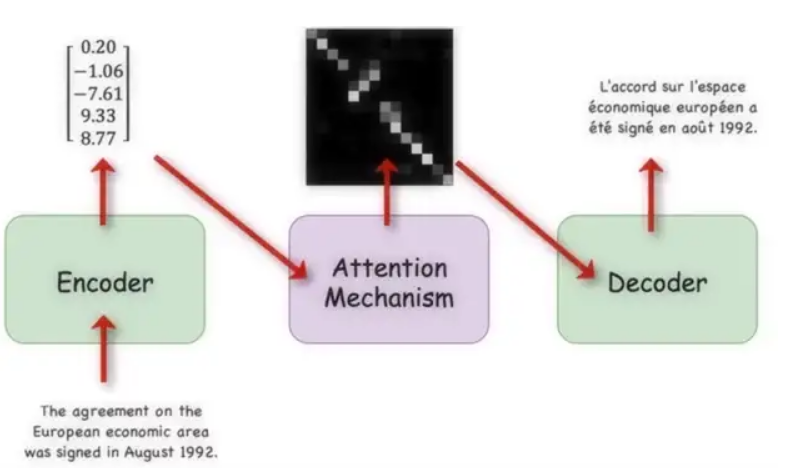
* metin içeren bir resim

  Açıklama otomatik olarak oluşturulduLanguage Translation: Components: We can break translation into two components: the individual units and the grammar:
* We need to encode sequences of words into vector spaces to perform computations in a neural net. Because the words also have meaningful sequences, a recurrent neural network is suitable for this task:

****The encoder-decoder architecture breaks down after about 20+ word sentences. Longer sentences illustrate the limitations of a single-directional encoder-decoder architecture. Because language consists of tokens and grammar, the problem with this model is it does not entirely address the complexity of the grammar.

Specifically, when translating the nth word in the source language, the RNN was considering only the 1st n-word in the source sentence, but grammatically, the meaning of a word depends on both the sequence of words **before and after** it in a sentence.

To avoid this problem, we use the bidirectional LSTM model. The bidirectional model allows us to enter the context of both past and future words to construct an accurate encoder output vector.

As a result of these stages, the following question may come to our mind, which word should we focus on in the sequences? We can learn which words in the source language to focus on by storing the previous outputs of the LSTM cells, then ranking them by the relevancy of each, and picking the word with the highest score. Below, you can see how this looks in the graphs: the resulting architecture then embeds this Attention mechanism in between the Encoder and Decoder.

If we briefly summarize the Neural Machine Translation steps:

1. The encoder takes each word in the source language and encodes it into vector space
2. These vector representations of words are then passed into an attention mechanism which determines which source words to focus on while generating some output for the desired language.
3. This output is passed through a decoder that turns the vector representations into the target language

**References**

1. https://levity.ai/blog/how-natural-language-processing-works
2. <https://localizejs.com/articles/natural-language-processing-nlp/>
3. <https://www.youtube.com/watch?v=fOvTtapxa9c>
4. <https://www.statmt.org/wpt05/mt-shared-task/>
5. <https://machinetranslate.org/statistical-machine-translation>
6. <https://phrase.com/blog/posts/neural-machine-translation/#definition-of-machine-translation>
7. <https://www.youtube.com/watch?v=AIpXjFwVdIE>
8. <https://omniscien.com/faq/what-is-rules-based-machine-translation/#:~:text=Rules%2DBased%20Machine%20Translation%20(RBMT)%20systems%20were%20the%20first,meanings%20depending%20on%20the%20context>.
9. <https://towardsdatascience.com/neural-machine-translation-15ecf6b0b>
10. https://www.wikipedia.org/