Human in the Loop Machine Learning for Language Translation

Laksh Advani
Department of Computer Science
University of Colorado Boulder
Laad8452@colorado.edu

Human Centered Machine Learning

Abstract

This paper deal with the problem of Machine Translation for complex and large texts in particular. We regularly use services like Google translate for short conversations and simple text translation, but what happens when we use it with complex texts like medical documents? We can see that the results are below par. My proposal deals with a novel way to Human in the Loop Machine Learning to boost our confidence and eventually move towards a crowdsourced Human in the Loop system.

1 Introduction and Motivation

In 2014 BMJ published a paper: Use of Google Translate in medical communication: evaluation of accuracy by Sumant Patil and Patrick Davies. This paper explores the use of google translate for a variety of languages in context of medical phrases.
The results were very surprising.

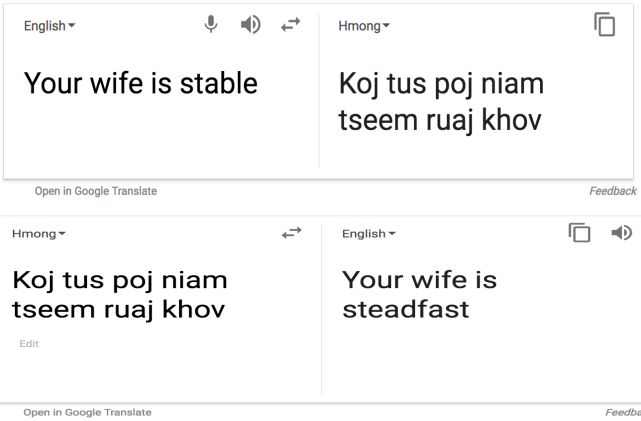
There were some serious errors. For instance, “Your child is fitting” translated in Swahili to “Your child is dead.” In Polish “Your husband has the opportunity to donate his organs” translated to “Your husband can donate his tools.” This demonstrates that common translation applications are still not universally accurate or usable for complex texts.

The paper concludes that Google Translate has only 57.7% accuracy when used for medical phrase translations and should not be trusted for important medical communications.

The motivation behind this problem is obvious as, the world becomes more interconnected the need for accurate translation is very important.

Assuming Google translate is the best translator application in production due to the sheer volume of data it has over competitors a 57.7% accuracy is an issue which we need to rectify by novel means.

Many people in low income and disaster affected areas of the world may need a better translation system to properly diagnose patients. Some results from the google assistant are below.



These examples have been taken from the google translate website.

Many microbial pathogenspathogenic microbes use adhesion receptors to penetrate mammalian cells. EnteropathogenicEntereropathogenic yersinia, for example, gain entrance intoenters the body from the lumen of the intestine,intestinal lumen, where they enter Peyer's patchesthe Peyer patch and cause mesenteric lymphadenitis. The invasin protein on theInvasin proteins in bacteria mediatesmediate the penetration of these gut structures; itthis bowel structure; This binds severalsome β1 integrins through ana series of amino acid sequenceacids containing an aspartic acid (as does thewell as an RGD tripeptide).triplet). Invasin also induces phagocytosis, as is demonstratedshown in vitro by the rapid uptake of invasin-coated microbeadsinvasive microbeole absorption coated by phagocytes.phagocytes rapidly. Characteristic signs of integrin signaling —- protein phosphorylation, alteration of cytoplasmic pH, and cytoskeleton rearrangement of the cytoskeleton — herald the- contain phagocytosis of the microorganism.microorganisms.

The figure above shows the additions(green) and deletions(red) in a medical translation from Indonesian to English. From these examples, we can see that there is still a lot for work to do. We can see that longer documents are typically poorly translated.

This problem is difficult for humans because it isn't practical for a human to have a working knowledge of say more than 10 languages at once, in many cases most humans know 2 languages. This is an issue when we need to collaborate or conduct medical tests.

We can also classify this problem as delegable to machines. This is because a perfect machine translation system can produce a very positive impact in the global community can solve the issue of translation which has been unsolved for centuries.

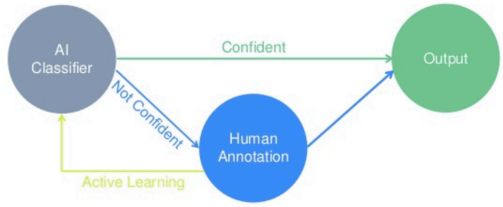
We can say that- this problem is delectable and difficult for humans. Human in the Loop Machine Learning is a promising methodology to follow to solve this problem.

2 Role of Humans in the Project

Human translations methods are typically considered paramount due to their experience with multiple languages and local dialects and expressions such that humans are used for translation at the highest levels (UN, diplomatic visits etc.). Every language has its slang and colloquialisms, and most of these reflect deeper truths about a culture, these are difficult for machines to extract.

Humans with specialized knowledge of the language demonstrate a much better performance against machines in translating languages based on previous experience. We can expect a

human translation to be perfect and can aid the machine to correlate the translations accurately. We can use a general Human in the Loop model to enhance accuracy, a team of target language medical experts can be assembled to help us with this.



A typical active learning workflow

Our team of humans will have two uses, they will have to handle cases of low confidence and will need to aid in the labeling of datasets. For lesser known languages like Hmong or Indonesian, as we saw in the examples above, the human team will need to classify the features of a text piece such as slang and colloquialisms.

In summary, the role of humans in this project is to generate a perfect translation to aid the machine learning process.

3 Role of Machines in the Project

Machine Translation has been around for quite some time, there are many existing services like Google Translate and Bing!. The scope for improvement lies with longer texts and accurate translations for complex articles.

The machine learning model I am proposing will be based on a neural network that uses a sequence to sequence model. It follows a general-purpose encoder-decoder framework for Tensor flow that can be used for Machine Translation. This is similar to Google's machine translation system, however ours will use a special register which allows us to integrate the human's input into the system. The sentences and words that we match with their non-English counterparts will be represented using n-grams which will have a set of weights. As the human input gets used, the weights for the correct translation increases and we begin to train our model.

In summary, the role of machines in this project will be to create an accurate classifier which will allow us to translate n-grams correctly.

4 Approach

Our work will initially begin with a corpus of English and our target language, take German as an example, we will first train our model to translate n-grams in this language. Next, we will need to analyze n-grams to begin to gain correlation between words and add them to our model's training system. Once we have a decent confidence score we will add the Human in the Loop segment, the introduction of a German medical language expert will help us classify difficult words or sets of words and update the weights for the n-grams in our special register.

As we accumulate human insights we should begin to see our model improve in accuracy and learn from the n-grams with updated weights form our special register.

5 Conclusion

In conclusion, this proposal may lead to an improved text translation service that can also work to join correlation over sets of n-grams or sentences. For the purpose of this class, this proposal is in line with Human Centered Machine Learning and is in the delegable and hard for humans' quadrant, this effectively makes this a high impact solution.

References

[1] Sumant Patil &Patrick Davies (2014) Use of Google Translate in medical communication: evaluation of accuracy In BMJ Christmas 2014.

[2] Ilya Sutskever, Oriol Vinyals, Quoc V. Le (2014) Sequence to Sequence Learning with Neural Networks from Google Whitepapers