Case Study NO. 025

Increasing quality and reducing time to market by utilizing a client-specific trained Machine Translation solution

How Neural Machine Translation (NMT) engines can be customized to tackle large material volumes and client-specific corpora to increase quality while saving time and costs.

UNITED LANGUAGE GROUP®

Background

One of ULG's largest clients needed a high volume of regulated material to be translated on a cyclical basis from English into seven languages, including: Hindi, Japanese, Korean, Russian, Spanish, Thai and Vietnamese.

The client required expedient turnaround and reliable, tested quality output to ensure the materials conveyed total accuracy for a multilingual audience.

The Challenge

After assessing the situation, it was clear several hurdles existed that would need to be overcome in order to successfully translate the high volume of materials and meet the critical deadlines of each cycle.

One difficulty was that the client-specific corpora needed to be accurately and quickly translated. Corpora can include anything from company logos to product names and lingo that are used solely by the client. Other in-market engines are incapable of deciphering not only the corpora, but the more complicated, non-traditional languages themselves, like Vietnamese.

To meet the turnaround requirements and ensure the quality of the output across languages and corpora, it was determined a customer-specific NMT engine would need to be created.

ULG's dedicated NMT team gathered all necessary industry and client specific content before completing the following steps: data cleaning and preparation, training and multiple levels of testing. These phases were finished before a customized engine could be released for utilization within the translation workflow.

The Solution

To create a customized engine a minimum amount of 15,000 clean segments is needed. Considering that Translation Memories (TMs) usually contain a significant amount of repetitions and segments that are just figures or unusable text, an amount of 30,000 segments from TMs are a good basis.

ULG's NMT team followed the standard engine customization process:

- Identify and extract eligible existing training material

- Perform a corpora cleaning process utilizing proprietary and third-party technologies

- Generate an acceptable set of clean and usable aligned segments to train the customized engine

The Results

ULG tested the BLEU scores of the in-market engines against the NMT engine that was custom built for the client. The results of the comparison illustrate large quality disparity between in-market and custom NMT engines.

In the below, the Live Production BLEU was able to reach the following scores:

Translation direction	Generic engine BLEU	Live production BLEU
English-Hindi	38.10	—
English-Japanese	35.52	46.89
English-Korean	38.12	43.50
English-Russian	20.88	55.67
English-Spanish	43.04	61.33
English-Thai	19.28	49.00
English-Vietnamese	31.63	57.00

The increase of the BLEU score was significant and allowed ULG to elevate the more difficult languages such as Russian and Thai to an appropriate quality level to utilize in production.

In conjunction with the quality-tested results, the customized NMT engine delivered significant benefits to the client including reduced turnaround time, and more than 38 percent reduction in time to market and overall savings.

What is a BLEU score?

ULG calculates the translation quality of its NMT engines using the standard BLEU quality score. This is a computer-generated metric that evaluates the quality of the results by comparing the NMT file to a reference file.

The in-market engines could provide an acceptable translation quality score for some languages for gisting purposes, but the output was too low of quality for even post-editing. Customized client-specific engines usually provide a translation quality of up to 10-15 BLEU score points (depending on the training data).

ULG uses the industry BLEU score as well as TER (automated edit distance scoring) and independent Distance Scoring performed by a team of linguists to determine and validate quality scoring of NMT.