ISSN: 3105-1073 (OnIine)
Journal of Al-Itqan for Human Sciences
available online at
https://iraqiqsa.org/Journal/index.php/AIJHS/index
مجلة الاتقان للعلوم الإنسانية تصدرها جمعية علوم القران العراقية

The Duality of Machine Translation: A Critical Analysis of Literal and Free Translation Approaches in AI Systems

M.M. Marwa Ghanem Ibrahim Hamoudi Al-Samarrai

Imam Al-Azam University College

م .م. مروة غانم ابراهيم حمودي السامرائي

كلية الامام الاعظم الجامعة

ازدواجية الترجمة الآلية: تحليل نقدي للترجمة الحرفية والترجمة الحرة

مناهج الترجمة في أنظمة الذكاء الاصطناعي

Abstract

The fast development of machine translation (MT) based on Artificial Intelligence (AI) as a system of neural machine translation (NMT) has transformed the realm of cross-linguistic communication. Such systems usually work within the range of literal (form-based) translation strategies and free (meaning-based) ones. The paper will critically discuss the inherent drawbacks of the literal and free translation methods as applied by AI. Based on a literature review, it contends that even literal AI translation is prone to generating grammatically aberrant, culturally insensitive, and semantically erroneous output, whereas free AI translation is prone to bring out uninformed distortion, loss of critical nuance, and bias. The paper concludes that the main problem is not the decision between the two extremes, but the fact that the AI is not yet capable of providing the complex contextual, cultural and pragmatic judgment in order to perform faithful and functional translations. The consequences to the users, developers and translation profession are addressed.

Keywords: **machine translation, AI translation, literal translation, free translation, neural machine translation, translation errors, cultural bias**

ملخص

أحدث التطور السريع للترجمة الآلية القائمة على الذكاء الاصطناعي، وتحديدًا نظام الترجمة الآلية العصبية، نقلة نوعية في مجال التواصل بين اللغات. تعمل هذه الأنظمة عادةً ضمن نطاق استراتيجيات الترجمة الحرفية (القائمة على الشكل) والترجمة الحرة (القائمة على المعنى). تتناول هذه الورقة البحثية بالنقد والتحليل العيوب الكامنة في أساليب الترجمة الحرفية والحرة عند تطبيقها بواسطة الذكاء الاصطناعي. واستنادًا إلى مراجعة الأدبيات، تُشير الورقة إلى أن الترجمة الحرفية للذكاء الاصطناعي عُرضة لإنتاج نصوص غير سليمة نحويًا، وغير مراعية للحساسيات الثقافية، ومُضللة دلاليًا، بينما تُعرض الترجمة الحرة للذكاء الاصطناعي لتشويه غير مدروس، وفقدان الفروق الدقيقة الهامة، والتحيز. وتخلص الورقة إلى أن المشكلة الرئيسية لا تكمن في الاختيار بين هذين النقيضين، بل في أن الذكاء الاصطناعي لا يزال غير قادر على توفير التقييم السياقي والثقافي والبراغماتي المعقد اللازم لإجراء ترجمات أمينة وفعّالة. وتتناول الورقة أيضًا التداعيات على المستخدمين والمطورين ومهنة الترجمة.

الكلمات المفتاحية: الترجمة الآلية، الترجمة بالذكاء الاصطناعي، الترجمة الحرفية، الترجمة الحرة، الترجمة الآلية العصبية، أخطاء الترجمة، التحيز الثقافي

Introduction

The application of Artificial Intelligence (AI) to machine translation (MT) has resulted in the highest level of availability in cross-linguistic communication. Neural machine translation (NMT) systems (whereby sentences are processed as a whole and trained on large corpora of parallel text), have been shown to achieve high levels of fluency compared to the previous rule-based or statistical models (Popel et al., 2020). Nevertheless, these achievements leave AI translation systems with a basic dilemma in translation theory the disagreement between literal (or formal equivalence) and free (or dynamic/functional equivalence) translation (Nida, 1964). Within its training data and architecture, AI does not make any conscious choice of strategy, but is statistically motivated to output within this space. It is the thesis of this paper that both inclinations, that is towards literalism and towards excessive liberalism, have their own unique and substantial downsides that undermine the quality of translation, the adherence to culture, and the communication effectiveness. The Cons of Literal Translation in Artificial Intelligence. The literal translation in AI happens when the system focuses on the superficial lexical and syntactic similarities between the target and the source languages. This is usually due to excessive dependence on direct pattern matching of its training cases or insufficient contextual awareness.

Grammatical and Syntactic Incoherence: Languages are syntactically different.

A literal translation of a subject-object-verb (SOV) language such as the Japanese into a subject-verb-object (SVO) language such as English may result in nonsensical word order. Although NMT is performing better in this aspect, AI cannot restructure sentences completely to generate output that is easy to read (Almahasees, 2021). Semantic and Collocational Errors: Words tend to be polysemous (have more than one meaning). An AI would literally interpret the phrase river bank to mean financial institution since the particular collocation is more common in its training data. On the same note, it can be poor at interpreting idioms (e.g., a literal translation of it rains cats and dogs), which will lead to confusion or ridiculousness (Bender, 2020). Linguistic: Cultural Insensitivity and Inaccuracy: Literal translation does not take account of the culture. Putting words on individual cultural ideas (e.g., in Danish: hygge) in word-to-phrase translation becomes deprived of its meaning. Worse still, the direct translation of an honorific, gendered language, or a culturally-specific item may result in distortion or insult. A Japanese refusal that was translated by an AI as polite it includes the words tiyotuto... / chotto... which translates to a little... this entirely lacks the practical purpose of civil disavowal.

The Negatives of Free Translation in AI

With a free translation, where the AI does not focus on the correspondence between the original meaning and the target language and focuses more on its overall meaning and naturalness in the target language, one opens a new collection of problems, most of which are issues of over-generalization and unclear algorithmic decision-making. Uncontrollable Distortion and Omission: In the quest to become fluent AI can simplify, leave subtleties out, or include explanatory information that does not exist in the original work. It is especially risky with legal, medical, or technical translation where accuracy is of the utmost importance. There could be a slight qualifier, a two-negatives or a particular technical word that the AI could smooth over, changing the intended meaning of the text and with potentially severe consequences (Forcada, 2017). Loss of Stylistic and Authorial Voice: Literary texts, marketing text, and academic writing have certain stylistic decisions to make. An AI that translates freely

could standardize the individual authorial voice to generic, fluent prose that eliminates rhetorical devices, metaphoric language, or purposeful ambiguity that are the focus of the text (Kenny, 2022). Examples Amplification of Bias and Ethical Risk: AI models are trained by human-created data which includes biases of society. Under the free translation style, the AI can actively make the text normalized based on these biases. As an example, a doctor in gramaciously gendered languages may be automatically gendered as he, a nurse may be gendered as she, and will replace culturally charged words in a manner that will propagate stereotypes (Prates et al., 2020). This implantation of prejudice is not usually noticeable to the final user, who gets a fluent yet ideologically biased text.

The Core Problem: Lack of Discernment

It is not that AI uses literal or free strategies, but the fact that it does not have the meta-linguistic and socio-cultural insight of the human translator to decide when to use either of them. Human translators are conscious and context-specific decisions making them use literalness when writing a patent specification, cultural adaptation when writing a marketing slogan, and a combination of both in a novel. AI is probabilistic, in the sense that it does not actually comprehend context, intention, or the communicative intent of a text (Bender and Koller, 2020). Its moves are not intentional but rather correlations and therefore its movements between being too literal and being too free translation are uncertain and outside context.

Conclusion and Implications

The drawbacks of literal and free translation in AI systems are used to underline the fact that fluency cannot be considered as fidelity or accuracy. This is the most important lesson that the user must be aware of the fact that AI translated output, whether it sounds strangely literal or alarmingly well-crafted, may contain semantic errors, cultural distortions, or implicit biases. In the case of developers, the difficulty lies in going beyond the quest to enhance fluency and create systems that are more context-aware and controllable to enable users to write translation briefs (e.g., "prioritize terminological consistency" vs. "adapt to a general audience). In the case

of the translation profession, such analysis confirms the inability of the human expert to be replaced. The best AI application would be in a human-in-the-loop approach, in which AI takes on first draft translation or high-volume, low risk content, and human translators use their judgment to revise, repairing the vices of both literalness and excessive freedom. Finally, it is important to comprehend the nature of AI inherent flaws in the literal-free spectrum to properly and ethically use it in our increasingly global world.

References

Almahasees, Z. (2021). Analysing English-Arabic machine translation: Google Translate, Microsoft Translator, and Sakhr. Machine Translation, 35(1), 29-44. https://doi.org/10.1007/s10590-021-09266-0

Bender, E. M. (2020). On achieving and evaluating language-agnostic NLP. Proceedings of the 4th Workshop on Research in Computational Linguistic Typology and Multilingual NLP (pp. 1-12). Association for Computational Linguistics.

Bender, E. M., & Koller, A. (2020). Climbing towards NLU: On meaning, form, and understanding in the age of data. Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (pp. 5185-5198). Association for Computational Linguistics.

Forcada, M. L. (2017). Making sense of neural machine translation. Translation Spaces, 6(2), 291-309. https://doi.org/10.1075/ts.6.2.06for

Kenny, D. (2022). Machine translation for everyone: Empowering users in the age of artificial intelligence. Language Science Press.

Nida, E. A. (1964). Toward a science of translating: With special reference to principles and procedures involved in Bible translating. Brill.

Popel, M., Tomkova, M., Tomek, J., Kaiser, Ł., Uszkoreit, J., Bojar, O., & Žabokrtský, Z. (2020). Transforming machine translation: A deep learning system reaches news translation quality comparable to human professionals. Nature Communications, 11(1), 1-15. https://doi.org/10.1038/s41467-020-18073-9

Prates, M. O., Avelar, P. H., & Lamb, L. C. (2020). Assessing gender bias in machine translation: A case study with Google Translate. Neural Computing and Applications, 32(10), 6363-6381. https://doi.org/10.1007/s00521-019-04144-6