Journal of Clinical Trials and Case Studies

Open access

A Comparison of BERT and T5 Models for Machine Translation: Evaluating the Performance of Google Translate and DeepL

Zaki Muhammad Zayyanu*

* *Department of , Faculty of Arts, Usmanu Danfodiyo University, Sokoto, Nigeria*

REVIEW ARTICLE

ABSTRACT

This paper is a continuation of an earlier paper titled Revolutionising Translation Technology: A Comparative Study of Variant Transformer Models – BERT, GPT and T5. It presents a comparative analysis of two leading transformer-based Natural Language Processing (NLP) models, BERT and T5, against prominent online machine translation services: DeepL and Google Translate. The researcher leverages on the theory of meaning and an interpretative approach. The research evaluates translation performance on a diverse set of ten (10) extracted datasets encompassing content from online newspapers, technical documents, and social media posts. The key metrics for assessment include translation accuracy, logical coherence, word order, and translator's experience. The findings reveal DeepL's superior accuracy in translations, particularly when integrated with BERT's language model, while also noting DeepL's User Interface (UI) advantage of a side-by-side text view and more comprehensive explanations. Google Translate is highlighted for its innovative features like "Word Lens", "Auto-detect", and accessible translation history. This comparative study not only examines the strengths and limitations of each service and tool but also discusses the implications of these findings for users and translators, ultimately leading to suggestions for future research in the domain of machine translation and NLP.

KEYWORDS

BERT, T5, Machine translation, Performance, Google translate, DeepL

Correspondence to: Dr. Zaki Muhammad Zayyanu, Department of , Faculty of Arts, Usmanu Danfodiyo University, Sokoto, Nigeria, Email: muhammad.zayyanu@udusok.edu.ng

Received: April 25, 2024, **Accepted**: May 06, 2024, **Published**: May 13, 2024

DOI: 10.1042/JCTCS.6.1.0013

Citation: Fatima Ezzahra Rizkou, 2024, A Comparison of BERT and T5 Models for Machine Translation: Evaluating the Performance of Google Translate and DeepL. JCTCS. Vol 6.

INTRODUCTION

Translation is an old practice in the history of humanity. It is an act of transforming a message from one language to another by maintaining the meaning of the original text. It is said to be a duplicate of the original. Later, Translation Studies surfaced as a field that deals with branches of pure and applied translation. Translation Technology has recently surfaced from the advancement of Machine Translation.

Machine Translation is the process of translating one language into another using Artificial Intelligence (AI). AI mimics the human brain and intelligence through neural networks to access different files within a second. This makes its speed unique as it can access several files within a limited time. Some experts in the field like Riu has discovered that Machine Translation has greatly improved translation efficiency and has become an integral part of scientific and cultural communication between different countries. However, Machine Translation still has limitations in conveying complex sentences and ideas, as it relies on literal translation and lacks the ability to understand context. Over the years, Machine Translation models have evolved from rule-based and statistical approaches to Neural Machine Translation (NMT), which has shown promising results, Zhao [1]. Machine Translation can replace human translation in certain areas, such as business and tourism. It cannot fully replicate the faithfulness, expressiveness, and elegance of human translation. This is influenced by social background and culture. Despite its limitations, Machine Translation is seen as a valuable tool for communication and interaction in the metaverse, with the potential to expand language availability and improve translation engine training.

Nevertheless, NMT is a new approach to Machine Translation that aims to build a single neural network capable of maximising translation performance. According to Dzmitry [2] and Thang [3], NMT models typically consist of an encoder-decoder architecture, where the encoder encodes a source sentence into a fixed-length vector, and the decoder generates a translation based on this vector. However, the use of a fixed-length vector can be a limitation in improving performance. To address this, recent research by Jiajun [4] and Alexandra [5] has proposed allowing the models to automatically search for relevant parts of the source sentence without explicitly forming them as hard segments. This approach has achieved comparable translation performance to existing state-of-the-art systems.

Moreover, Translation models are used for various purposes such as training methods, improving automatic speech recognition, and simulating non-stationary processes. These models involve encoding and decoding layers and can be trained using parallel text or historical translation texts. They can be used to obtain translated texts from input texts, improving response speed and user experience. Additionally, according to Oli[illegible] [6], Tianchi et al., [7] and Changliang, [8] translation models can be used to learn lexical translation parameters directly from word frameworks, resulting in word error rate reductions. Optimisation algorithms can also be applied to construct translation models that match the marginal distribution and covariance function of Gaussian processes, allowing for their use in Monte Carlo simulation studies, Ferrante et al., [9] and [illegible]orewaard., H., et al [10].

Also, Bidirectional Encoder Representations from Transformers (BERT) and Text-To-Text Transfer Transformer (T5) models are both Large Language Models (LLMs) that have been investigated for various Natural Language Processing tasks. In the context of email spam detection, LLMs, including BERT-like models, Sentence Transformers, and Seq2Seq models, have shown superior performance compared to traditional machine learning techniques such as Naive Bayes and Light GBM, [11-14]. Additionally, a specific adaptation of the T5 model called Spam-T5 has been introduced and fine-tuned for email spam detection, outperforming baseline models and other LLMs, especially in scenarios with limited training samples [15]. In the field of relation extraction for drug-protein interactions, BERT-based models and T5 models have been explored. While larger BERT-based models generally perform better, the T5 text-to-text approach shows promising results and has potential for further research advancement. T5 models have also been investigated for producing sentence embeddings, with encoder-only models outperforming BERT-based sentence embeddings on transfer tasks and semantic textual similarity (STS). Finally, an ensemble model combining BERT, sentence BERT, and T5 models achieved high performance in predicting drug-protein interactions based on sentence semantics.

The way the BERT model operates is by pre-training a language model in a bidirectional fashion on a sizable corpus of text. It allows the model to understand context from both the left and the right sides of a word. The BERT's strength is on tasks like text completion and question answering the demand through comprehension of the context [illegible] BERT can capture complex interactions that exist between words in a sentence because of its bidirectional approach. Its drawback is that it needs a lot of data and processing power to train. It is slower for longer texts and requires a lot of computing power because it sequentially processes text. While T5 is a flexible transformer-based concept that Google Research unveiled in 2019. T5 unifies many tasks under a common text-based architecture by framing all NLP tasks as text-to-text tasks, in contrast with BERT. It can perform a variety of tasks because it has been pre-trained to translate between different text formats. T5's task-specific designs are made simpler because of the text-to-text framework's strength gives it great flexibility and ease of application for a wide range of NLP tasks, Zaki. T5 achieves state-of-the-art performance across a variety of benchmarks because of its unified architecture. The training data's quality and variety for a wide range of tasks may have influenced T5's performance constraints. T5 model training and optimisation can need a lot of resources, particularly for large-scale applications.

DeepL and Google Translate (GT) are machine translation systems that have been used in various studies. GT and DeepL have shown improvements in their machine translation results, leading to their increased use by students and translators [16]. These systems have been evaluated in terms of their performance in translating specialised texts, such as legal arbitration clauses, and specialised lexical combinations. The results indicate that while the translations generated by GT and DeepL are more "fluent" than "adequate," they still have shortcomings in terms of terminological accuracy and handling of phraseological variation and discontinuity [17]. Despite these limitations, the translations of specialised lexical combinations were considered satisfactory [18]. Overall, GT and DeepL are popular neural machine translation systems that have shown promise in certain translation tasks, but they still face challenges in handling specialised texts and maintaining terminological consistency.

The research objectives are:

- Can these transformer models improve translation performance?
- Can these transformer models aid Machine Translation to increase translation quality?

The researcher will respond to these in the quest of this research. With the advancement in translation technology, most translators are not familiar with some models integrated into some translation tools like DeepL and GT to produce effective results. Hence, the translator must know about translation technology to achieve consistency and speed in translating huge tasks.

LITERATURE REVIEW

BERT and T5 are both state-of-the-art Natural Language Processing (NLP) models based on the transformer architecture. While they share some similarities, they also have distinct characteristics. Some similarities between BERT and T5 are: Both BERT and T5 are built on the transformer architecture introduced by Vaswani et al. in the paper "Attention is All You Need". This architecture allows for efficient parallel processing of sequential data, making it well-suited for NLP tasks.

BERT and T5 both follow a two-step process in which they are first pre-trained on a large corpus of text data and then fine-tuned for specific downstream tasks. Pre-training involves learning contextualised representations of words or subwords, capturing the relationships between words in a sentence. Both models benefit from large-scale training on massive datasets, enabling them to capture intricate patterns and nuances in language. BERT, for example, is pre-trained on a massive corpus such as the Book Corpus and English Wikipedia, while T5 is trained on a similarly extensive dataset.

BERT and T5 leverage transfer learning to improve performance on downstream tasks. The pre-trained models serve as strong feature extractors, and the knowledge gained during pre-training is transferred to specific tasks through fine-tuning. Both models use subword tokenisation, which involves breaking down words into smaller units (sub words or tokens). This enables the models to handle a larger vocabulary and capture morphological variations in languages effectively.

Both BERT and T5 utilise a **Masked Language Model (MLM)** objective during pre-training. In BERT, random words are masked, and the model is trained to predict the masked words based on the surrounding context. T5 extends this idea to a text-to-text framework, where the input is converted into a text template, and the model is trained to generate the output.

Task Agnostic Representations

Both models learn task-agnostic representations during pre-training, meaning that the pre-trained models can be fine-tuned for a wide range of NLP tasks without task-specific modifications to the architecture.

A comparison between BERT and T5 models is made based on selected characteristics in this research which are accuracy, speed and translators' experiences. BERT is known for its strong performance in various Natural Language Processing (NLP) tasks, achieving state-of-the-art results in tasks such as question answering, text classification, and Named Entity Recognition (NER). Its bidirectional attention mechanism enables it to capture contextual information effectively. While T5 follows a text-to-text framework, treating all NLP tasks as a text generation problem. It achieved competitive results across multiple tasks and achieved state-of-the-art performance in certain benchmarks. T5's versatility lies in its ability to frame diverse tasks in a unified format.

BERT's bidirectional nature poses a challenge for parallelisation during training, making it computationally intensive and slower compared to models with unidirectional attention. However, for inference, there are optimised versions and techniques, such as quantisation and model compression, to improve speed. While T5's training time can vary depending on the size of the model and the dataset. The text-to-text framework allows for more parallelisation during training, which can contribute to faster training times compared to BERT in certain scenarios. Inference speed can be optimised through model compression and deployment strategies.

BERT may require task-specific fine-tuning for optimal performance on a particular NLP task. It has been widely adopted and has a mature ecosystem of pre-trained models for various tasks. Translators may need to have experience with fine-tuning BERT for specific translation tasks. While T5's text-to-text framework simplifies the approach to various NLP tasks, making it easier for translators to work with a unified model. The translation task, for example, can be framed as a text generation problem where the source language is the input text, and the target language is the generated text. This can potentially streamline the translation process. It is important to note that the choice between BERT and T5 depends on the specific requirements of the task at hand, available computational resources, and the characteristics of the data.

An overview is that DeepL and Google Translate are popular machine translation services, each with its own strengths and weaknesses. On one hand, DeepL is a Neural Machine Translation (NMT) service developed by the German company DeepL GmbH. It gained attention for its ability to provide high-quality translations, often outperforming other machine translation systems. DeepL uses deep learning techniques and a large neural network to generate translations. **The key feature is** High Translation Quality. DeepL is known for providing translations that are often more fluent and contextually accurate than many other translation services. Then, DeepL supports translation between multiple languages, including English, German, French, Spanish, Italian, Dutch, Polish, and more. DeepL offers an API that developers can integrate into their applications for automated translation.

On the other hand, Google Translate is a machine translation service developed by Google. It has been available for many years and is widely used globally for translating text and websites. Google Translate uses statistical machine translation and, more recently, has incorporated neural machine translation for certain language pairs. Its key f**eatures are:** Google Translate that supports a vast number of languages, making it one of the most widely used translation tools globally. While the online version requires an internet connection, Google Translate also offers offline language packs for certain languages. Google Translate is integrated into various Google services, such as Chrome browser and Google Docs.

The comparison between DeepL and Google Translate in this research are quality, language support, Application Programming Interface (API) API access and Translators' interface. DeepL is often praised for its high translation quality, which is attributed to its neural network architecture. While Google Translate through widely used, may not always match DeepL in terms of quality. DeepL makes it a more versatile choice for users who need translation in less common languages. While Google Translate supports a larger number of languages compared to DeepL. Both DeepL and Google Translate offer API access, allowing developers to integrate translation services into their applications. And, DeepL provides a clean and simple interface as well. While Google Translate has a user-friendly interface and is easily accessible online. A screenshot of the layout is presented below:

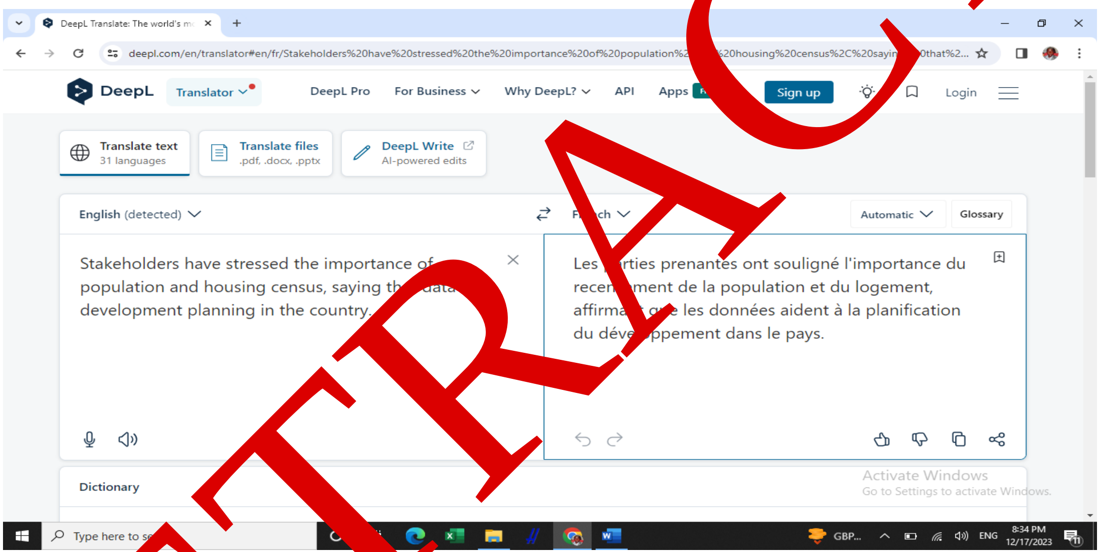

Figure 1: DeepL environment.

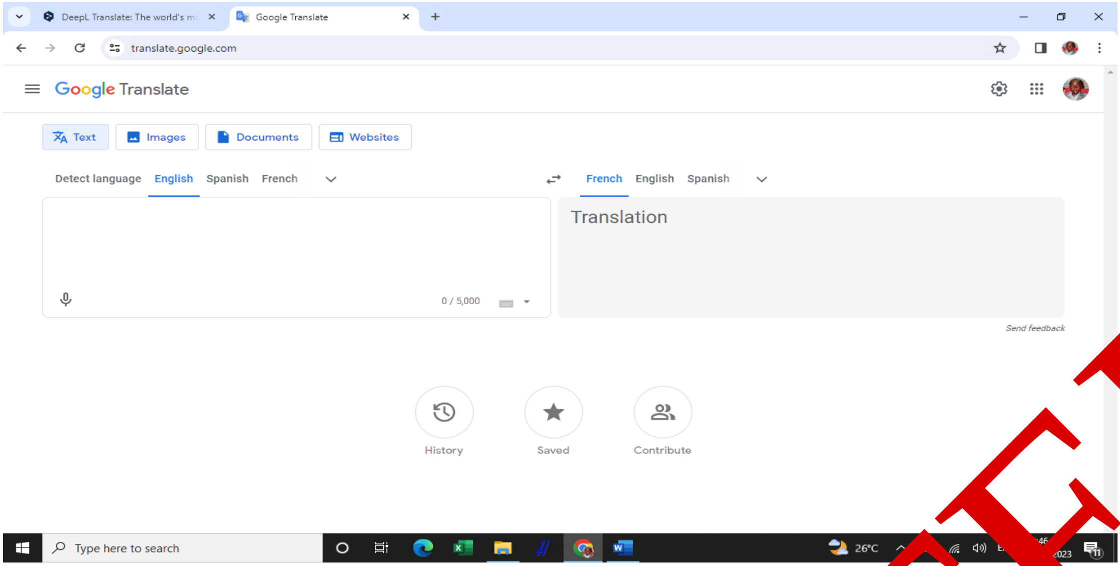

Figure 2: Google Translate environment.

METHODOLOGY

The researcher applies the theory of meaning and interpretative approach in the evaluation. Datasets are a good way of evaluating the tools. The researcher extracted datasets that contain a variety of text types. These datasets are extracted from online newspapers, technical documents, and social media posts. This allows the researcher to test the tools in different scenarios and get a comprehensive evaluation. The researcher will evaluate the dataset by testing the translated texts, word order, logical presentation and the accuracy of each translation tool (DeepL and Google Translate).

DATA ANALYSIS

The researcher presented extracted datasets from online newspapers, technical documents and Social media posts. The datasets are four (4) extracts each for some, making a total of ten (10) extracts. The datasets are arranged accordingly from English to French, two (2) extracts each totaling four (4) extracts and from French to English, two (2) extracts each totaling six (6) extracts. These extracted data sets are made up of ten (10) extracts to evaluate the performance of the tools, especially accuracy, they are also interpreted using the theory of meaning from the analysis of their performance. The analysis is presented thus:

English to French

Dataset extract 1:

Dataset	DeepL	Google Translate
Stakeholders have stressed the importance of population and housing census, saying that it aids development planning in the country.	Les parties prenantes ont souligné l'importance du recensement de la population et du logement, affirmant que les données **aident** à la planification du développement dans le pays.	Les parties prenantes ont souligné l'importance du recensement de la population et du logement, affirmant que les données **facilitent** la planification du développement dans le pays.

Table 1: Leadership page 2, (12th December, 2023).

Overall, both translations are accurate and convey the meaning of the original text. However, there are some subtle differences between the two. DeepL "aident" this is a more formal word choice that is more likely to be used in a written document.while GT "facilitent" this is a more literal translation that is not as idiomatic in French.

Dataset extract 2:

Dataset	DeepL	Google Translate
Just recently, at the 11th annual population lecture series 2023, organised by National Population Commission with the theme: “A decade of dialogue on population and development in Nigeria” in Abuja.	Tout récemment, lors de la 11e série **de conférences annuelles** sur la population 2023, organisée par la Commission nationale de la population sur le thème « Une décennie de dialogue sur la population et le développement au Nigeria » à Abuja**.**	Tout récemment, lors de la 11e série **annuelle de conférences** sur la population 2023, organisée par la Commission nationale de la population sur le thème « Une décennie de dialogue sur la population et le développement au Nigeria » à Abuja.

Table 2: Leadership page 2, (12th December, 2023).

The English extract “annual population lecture” is translated using DeepL application as “de conférences annuelles” while GT “annuelle de conférences” in French.

Dataset extract 3:

Dataset	DeepL	Google Translate
In a survey of 2,000 British adults, carried out by Vision Direct, it revealed that we spend 4,866 hours a year staring at screens, whether that be phones, computers, televisions, gaming devices or e-readers.	Une enquête **menée** par Vision Direct auprès de 2 000 adultes britanniques a révélé que nous passons 4 866 heures par an à regarder des écrans, qu'il s'agisse de téléphones, d'ordinateurs, de téléviseurs, **d'appareils de jeu ou de lecteurs électroniques**.	Une enquête **réalisée** par Vision Direct auprès de 2 000 adultes britanniques a révélé que nous passons 4 866 heures par an à regarder des écrans, qu'il s'agisse de téléphones, d'ordinateurs, de téléviseurs, d'appareils **de jeux ou de liseuses électroniques**.

Table 3: Leadership page 25, (12th December, 2023).

The English extract “survey” is translated using DeepL application as “menée” and GT translated as “réalisée” in French. And, another English extract « computers, televisions, gaming devices or e-readers” using DeepL application as “de jeu ou de lecteurs électroniques” while GT “ de jeux ou de liseuses électronique” in French.

Dataset extract 4:

Dataset	DeepL	Google Translate
The use of these devices can also lead to dry eyes and eyestrain.	L'utilisation de ces appareils peut également entraîner une sécheresse oculaire et une fatigue **visuelle**.	L'utilisation de ces appareils peut également entraîner une sécheresse oculaire et une fatigue **oculaire**.

Table 4: Leadership page 25, (12th December, 2023).

The two translations are okay, but with some room for adjustment. The English extract “eyestrain” is translated using DeepL application as “visuelle” while GT translated as “oculaire”. This is a good example of how both DeepL and Google Translate can provide accurate translations. The choice between them may come down to stylistic preference for hyphenation or rephrasing of certain words.

French to English

Dataset extract 5:

Dataset	DeepL	Google Translate
Pour comprendre l'embrasement du Mali depuis 2012, la journaliste Nathalie Prévost donne la parole à ses protagonistes. En 2022, dans un contexte sécuritaire dégradé par la détérioration de la relation entre Paris et Bamako, elle a parcouru le pays à la rencontre des acteurs maliens de la crise.	To understand the conflagration in Mali since 2012, journalist Nathalie Prévost gives a voice to its protagonists. In 2022, in a security context **worsened** by the deterioration of relations between Paris and Bamako, she traveled the country to meet the Malian **players** in the crisis.	To understand the conflagration in Mali since 2012, journalist Nathalie Prévost gives a voice to her protagonists. In 2022, in a security context **degraded** by the deterioration of the relationship between Paris and Bamako, she traveled the country to meet the Malian **actors** in the crisis.

Table 5: Nathalie Prévost, 19th December 2023.

Both DeepL and Google Translate provide accurate translations of the passage, but there are slight nuances in phrasing: DeepL "worsened, players" while GT is "degraded, actors". This revision uses "turmoil" as a more nuanced alternative to "conflagration". It replaces "protagonists" with "key players" for a more neutral tone. Both translations are good, but with some room for improvement. Here is a suggested revision. Finally, it clarifies "degrade" (degraded) with "worsening" and "acteurs" (actors) with "actors involved" for better clarity.

Dataset extract 6:

Dataset	DeepL	Google Translate
De Bamako à Kidal, en passant par Tombouctou, Mopti ou Anéfis, RFI vous propose les témoignages rares des leaders politiques, militaires et religieux maliens qui portent des espoirs et des visions contrastées de l'avenir de leur pays.	From Bamako to Kidal, passing through Timbuktu, Mopti and Anéfis, RFI **brings you** rare testimonies from Malian political, military and religious leaders, who carry their **contrasting hopes and visions of their country's future**.	From Bamako to Kidal, via Timbuktu, Mopti or Anéfis, RFI **offers you** rare testimonies from Malian political, military and religious leaders who carry **hopes and contrasting visions of the future** of their country.

Table 6: Nathalie Prévost, 19th December 2023.

The two translation software, DeepL and Google Translate provide accurate translations of the extracts, but there are slight distinctions in arrangement. DeepL translates "vous propose" in English as "brings you" while GT "offers you". Also, the French phrase « portent des espoirs et des visions contrastées de l'avenir de leur pays » translated by DeepL version as "contrasting hopes and visions of their country's future" and GT as "hopes and contrasting visions of the future".

Dataset extract 7:

Dataset	DeepL	Google Translate
Voici à quoi ressemble le boîtier Smart TV. C'est un petit boîtier qui se branche sur le port HDMI de la télé.	This is what the Smart TV box looks like. It's a small box that plugs into **the TV's HDMI port**.	This is what the Smart TV box looks like. It's a small box that plugs into the **HDMI port on the TV.**

Table 7: www.rfi.fr 19th December 2023.

The French extract "le port HDMI de la télé" is translated using the two applications: DeepL "the TV's HDMI port" while GT "HDMI port on the TV". The applications provide accurate translation but there are some differences. Both translations are good, but with some room for improvement.

Dataset extract 8:

Dataset	DeepL	Google Translate
L'avantage c'est que **toutes les télés à travers le monde sont équipées d'un port HDMI depuis 2003, donc tout le monde ou presque peut l'utiliser.**	The **good thing** is that all TVs **worldwide** have been equipped with an HDMI port since 2003, so almost everyone can use it.	The **advantage** is that all TVs **around the world** have been equipped with an HDMI port since 2003, so almost everyone can use it.

Table 8: www.rfi.fr 19th December 2023.

The two versions of extracts fron French "avantage" DeepL "good thing," while GT "advantage". On the other hand, "tout le monde" using DeepL is translated as "worldwide" while GT translated it as "around the world". Both translations are good, but with some room for improvement.

Dataset extract 9:

Dataset	DeepL	Google Translate
Une entreprise Canadienne vient de sortir un chauffage à très faible consommation d'énergie permettant de chauffer une maison à plus de 30°C quasiment gratuitement. Une révolution qui cartonne au Canada enfin disponible en France.	A Canadian company has just **launched an ultra-low-energy heating system** that can heat a house to over 30°C **virtually free of charge**. A revolution that's been a hit in Canada, now available in France.	A Canadian company has just **released a heater with very low energy consumption** that can heat a house to more than 30°C **almost for free**. A revolution that is a hit in Canada, finally available in France.

Table 9: Julien Lacroix - Rédacteur chez Astuce Chauffage 19th December 2023.

The French version of the extract "**de sortir un chauffage à très faible consommation d'énergie" is transformed using the applications.** The two translations using DeepL is translated as "launched an ultra-low-energy heating system" while GT "released a heater with very low energy consumption". And also, another extract "**quasiment gratuitement**" is translated using DeepL as "virtually free of charge" while GT translates it as "almost for free".

Dataset extract 10:

Dataset	DeepL	Google Translate
Voici l'appareil Thermaly. Un chauffage dernière génération équipée de nanoéléments lui permettant de **consommer 99,8% moins d'énergie qu'un chauffage classique.**	**Introducing** the Thermaly heater. A latest-generation heater equipped with nano-elements **that consume** 99.8% less energy than a conventional heater.	**Here is** the Thermaly device. A latest generation heater equipped with nanoelements **allowing** it to consume 99.8% less energy than conventional heating.

Table 10: Julien Lacroix - Rédacteur chez Astuce Chauffage 19th December 2023.

The French extract "appareil" is translated using DeepL translation application as "Introducing" while GT translated it as "here is". Also, DeepL translate "lui permettant" as "that consume" and GT as "allowing". Both translations are good, but with some room for improvement. **Overall, the researcher would say that DeepL is the better translation of this text.** It is more accurate, more idiomatic, and uses more formal language. However, Google Translate is still a good translation and may be more suitable for some users.

DISCUSSION OF RESEARCH FINDINGS

BERT and T5 are language models that have proven to be very effective for machine translation. They are trained on large amounts of text data which allows them to learn the structure and relationships between words in a sentence. In this research, DeepL and Google Translate tools are evaluated for their performance in accuracy, speed and translator's experience. They use BERT and T5 or its equivalence for some translation tasks if integrated.

On one hand, DeepL is a great tool in translation and even better when used in combination with BERT. It uses advanced settings that can be used to improve translation quality and accuracy when selected. The option is located at the top of the page when using the tool. It is labeled as "options". Open the menu click on "Customise Translations" and choose DeepL's own language model (BERT). Then, select the BERT option for the best result.

Word order in both DeepL and GT may be the same. However, it is important to note the quality of the translation may vary between the two tools. GT may be more likely to make errors in word order than DeepL. This is because translation tools are not perfect, and there can be differences in the quality of translations between the tools. However, what the researcher observed is that these tools are constantly improving. Google Translate, for example, has been making significant improvements in recent years. These improvements include the integration of BERT and T5 translation models. Integrating BERT in GT has options that are selected to improve the quality of translation. The option is called "Enable Neural Machine Translation". This option uses a neural network to improve the quality of the translations. It is not quite as powerful as BERT but it is still a useful option in GT. Equally, T5 is another language model similar to BERT. GT has T5 available in its options. The translator enables it in "options" and scroll from the menu and selects "T5". Just like BERT, T5 can improve the quality of the translation.

On the other hand, GT uses BERT to improve the accuracy and fluency of its translation. It has open-sourced its BERT model making it available for use. GT uses the BERT model to make translation more accurate and naturally sounding. It is discovered that when concise and simple language is used in source texts, it makes it easier for BERT to understand the text and produce a more accurate translation.

From the research findings, the researcher found out that it is not possible to use T5 in DeepL while it is possible in GT. There are some differences between the two tools. DeepL allows translators to see the original text and the translation side-by-side while GT does not. In addition, DeepL provides more detailed explanations of the translation while GT explanations are more general. The researcher takes into account the different features of each tool. DeepL has a built-in grammar checker while Google Translate does not. This is a limitation that can be addressed through the evaluation of each tool feature.

DeepL has also useful features like "Dictionary" that allows translators to look up the meaning of words. It has a "phrasebook" that allows a translator to save frequently used phrases. It has a "profanity filter" that allows translators to remove profanity from translations. GT has many features that make it a powerful transition tool. It has a "Word Lens" that allows translators to translate text by pointing their camera at it. It also has an "Auto-detect" that automatically detects the language of the input text. It has a "History" that allows translators to view their translation history.

There are common features to both DeepL and GT. Both tools allow translators to select source and target languages. They both allow translators to copy and paste texts for translation. They also have a common tool that allows translators to listen to the translated texts aloud.

CONCLUSIONS

This research presents a comprehensive evaluation of two leading Machine Translation services, DeepL and Google Translate, in the context of their use of state-of-the-art NLP models BERT and T5. Our findings underscore DeepL's superior translation accuracy, particularly with specialised texts, which might be attributed to its advanced settings that prioritise translation quality. Despite this, Google Translate delivers commendable performance with the advantage of supporting a broader range of languages and the integration of advanced NLP models to enhance accuracy and fluency, although without features like side-by-side text comparison and a built-in grammar checker that DeepL offers. BERT and T5 have revolutionised tasks of Machine Translation. Both services benefit from the advancements in NLP, as evidenced by the use of transformer-based architectures of BERT and T5. Google Translate's incorporation of these models is an encouraging sign that Machine Translation services are evolving to leverage cutting-edge research for practical applications.

Nevertheless, the analysis reveals limitations rooted in focusing solely on DeepL and Google Translate while other significant translation tools remain unexamined. This gap suggests the need for further research that includes a wider array of translation services, as well as a deeper exploration into the potential application of the BERT and T5 models in services of DeepL to determine its impact on translation quality.

Eventually, as language models continue to advance and Machine Translation services refine their algorithms and interfaces. The prospect is promising for both linguistic accessibility across the globe and advancements in the quality of automated translations. Users stand to benefit from this competitive market as companies strive for higher accuracy, speed, and user experience in removing language barriers.

RESEARCH LIMITATIONS AND FUTURE DIRECTION

The limitations in language coverage and translation of less common language pairs by models of BERT and T5, as well as translation services of DeepL and Google Translate need to be addressed. Future research should aim to increase linguistic diversity and explore the scalability of these models to support more languages, specifically low-resource ones.

REFERENCES checd

1) Zhao L. The Relationship between machine translation and human translation under the influence of artificial intelligence machine translation. Mobil Inform Sys. 2022, 23:2-9.
2) Dzmitry B, Cho K, Bengio Y. Neural machine translation by jointly learning to align and translate. Comp Lang. 201[illegible].
3) Thang L. Neural machine translation. 2016.
4) Jiajun, Z, Zong CQ. Neural machine translation: challenges, progress and future. Science China-technological Sciences. 2020, 63:2028-2050.
5) Alexandra B. Neural Machine Translation 2020, by Philipp Koehn, Cambridge, Cambridge University Press, pages 393. Natural Language Engineering. 2021.
6) Oliver A, Neubig G, Cohn T, Bird S. Learning a translation model from word lattices. Conference: Interspeech. 20[illegible].
7) Tianchi B. Translation model-based translation method and device and storage medium. 2019.
8) Changliang L. Translation model training method and device. 2020.
9) Ferrante FJ, Arwade SR, Graham-Brady LL. A translation model for non-stationary, non-Gaussian random processes. Probabil Eng Mechanics. 2005, 20:215-228.
10) Doorewaard H, Knudsen T. Translation theory and the integreted approach in it management. [illegible]DC 2000 Conference Proceedings. 2000.
11) Adams V, Hoo-Chang Shin, Anderson C, Bo Liu, Abidin A. Text mining drug/chemical-protein interactions using an ensemble of BERT and T5 Based Models. Comp Lang. 2021.
12) Jianmo Ni, Abrego GH, Yang Y, Gustavo JN. Sentence-T5: Scalable sentence encoders from pre-trained text-to-text models. Comp Lang. 2021, 1864-1874.
13) Xin S. Text Mining Drug-Protein Interactions using an Ensemble of BERT, Sentence. 2021.
14) Judit Á, Lévai D, Kornai A. Evaluating transferability of BERT models on Uralic languages. 2021.
15) Sui X, Wang W, Zhang J. Text mining drug-protein interactions using an ensemble of BERT, sentence BERT and T5 models. In: Proceedings of the BioCreative VII Challenge evaluation workshop. 2021, 1:54-58.
16) Françoise M. DeepL et Google Translate face à l'ambiguïté phraséologique. J Data Mining Dig Hum. 2022, 1:1-16.
17) Carlos M. Google Translate vs. DeepL: analysing neural machine translation performance under the challenge of phraseological variation. MonTI - Monografías de Traducción e Interpretación. 2021.
18) Xuedan D, Cai Y, Wang S, Zhang L. Overview of deep learning. Youth Academic Annual Conference of Chinese Association of Automation (YAC). 2016.