

Error Analysis of Automatic Machine Translation on @Ajikdewa_43 Instagram Caption Posts

I Gede Arista Pramana Putra¹, I Gusti Ayu Mahatma Agung²
gdaris04@gmail.com¹, ayu.mahatma@unmas.ac.id²

Universitas Mahasaraswati Denpasar

Abstract

Instagram, one of the most popular social media platforms, is widely used by Indonesian people to connect and communicate with each other worldwide. The impact of Instagram influenced many economic sectors, such as marketing, creative industry, and personal branding. Therefore, social media is utilized through content production to gain additional income and extend engagement. One of the most commonly known Indonesian Instagram accounts is AjikDewa_43, which focuses on Balinese entertainment and non-entertainment news. Most of the posts are wrapped with general Balinese humor content supported by entertaining captions. However, due to systematic error, the automatic translation feature cannot correctly render the text into the target language of certain users. Therefore, this research aims to identify, categorize, and evaluate errors of automatic machine translation on Instagram caption posts with a case study of @Ajikdewa_43 media platform. Qualitative method with descriptive analytic approach was used to analyze the data. The data were obtained from 10 selected posts with short or long captions. Koponen's theory of error category was applied in this study to describe the translation quality of Instagram captions. The result showed significant errors of the machine translation in identifying slang and local language as well as the contextual and grammatical meaning of the language. Several identified translation errors linked with 4 basic translation concepts were found, namely Omitted Concept (10%), Untranslated Concept (40%), Mistranslated Concept (40%), and Substituted Concept (10%).

Keywords: Instagram, machine translation, translation error

Introduction

Instagram is widely known as a social media platform that connects worldwide users through content production such as video reels, original posts, and stories. The usage of Instagram in modern times has various needs. The most relevant activities nowadays are promoting digital marketing campaigns or entertainment purposes. People can upload content to its service and share it with their followers or a select group of friends (Anisah et al., 2021; Ting et al., 2015). According to Monthly Active Users (MAU), alongside other media platforms such as Facebook, YouTube, and WhatsApp, Instagram holds more than 2 billion active users worldwide (Rizaty, 2022). One of the countries that has a high number of active accounts is Indonesia, with over 100 million users.

The use of Instagram has drawn significant attention from researchers, leading them to conduct investigations related to this particular social media site. Astuti et al. (2020) examine a study of impulse buying by Instagram users in Indonesia and show indications that lead younger citizens to do online shopping. Three factors that support this habit are hedonic shopping motives, online advertisement content, and electronic word of mouth (E-WOM), commonly known as speech craft mastery. In terms of social media marketing, Instagram enables the success of activating business campaigns by allowing marketers to reach a broader

range of customers domestically and internationally. The tools provided by Instagram facilitate them in building relationships between businesses through online interactions (Leung et al., 2022). In addition, another research conducted by Darmaningrat et al. (2020) studied how Indonesian Instagram usage outside business purposes is viewed from privacy concerns among college students. They trust Instagram's reputation and the ease of using the platform while ignoring the risk of information misuse by using Instagram and instead continue socializing through available content.

In summary, both research regarding media policy and marketing communication informed us that most Indonesians use Instagram for online shopping through content production from one of the users, regardless of the information provided inside. This might raise awareness for Indonesian netizens that every user or account has to improvise their content production and segment their audience in order to appeal to the target market properly (Natasha et al., 2023). There are many active social platforms that aim for a specific audience to consume the posted content. Some of them are based on language and demography. One of the most entertaining media that shows common information and entertainment concerning Indonesia is AjikDewa_43.

The mentioned Instagram account provides humorous content, informational posts, and endorsements specifically for Balinese Instagram users because the post captions are primarily written in Balinese. However, other users cannot experience the communication originating from the account by only using one particular language. A similar Instagram account to AjikDewa_43, namely LawakBaline produces similar content, but they generally use Indonesian language in order to broaden the audience segmentation. An example of a case study of Folkative Instagram account by Adawiyah et al. (2023) shows several successful translations from English as the source language to Indonesian as the target language. Consequently, when they utilize the machine translation feature, it successfully renders the whole caption in one piece. Therefore, the machine translation allows them to understand the source language used in the caption.

In order to acknowledge the unfamiliar language, Instagram has a feature that automatically translates the source language of the caption to the target language of the general users. However, there are many mismatches and errors, such as unregistered and unsuccessful renders, due to the poor quality of machine translation. This occurs due to the fact that Instagram only provides general language and does not have a specific language (Nadilla & Lubis, 2023; Swadeshi & Sutrisno, 2021). Furthermore, translation machine on Instagram does not properly choose the best translation technique for certain part of the text, but rather generates a common word or sentence (Antika, 2022; Meilasari, 2019; Putri & Setiajidi, 2021). This happens because of the random analysis done by translation machine that does not clearly analyze what is being translated. Another research conducted by Kendenan et al. (2023) explains that every Indonesian traditional language, such as Javanese language, has different characteristics based on the politeness level of the language. Consequently, the translation machine could not convert small details of linguistic patterns from a certain language. Therefore, to identify several translation mismatches, this research is conducted to identify several translation errors in the AjikDewa_43 Instagram account, focusing on 10 post captions as the main data.

Theory and Method

This research used qualitative method with descriptive analytic approach. According to Faruk (2012), qualitative method is applied to describe the data contributed to the research and to obtain descriptive results. The primary data of this study was collected from AjikDewa_43 Instagram account. There were four steps in collecting the data, namely selecting 10 posts on AjikDewa_43 Instagram account that use Indonesian or Balinese language, translating the

caption using the “see translation” feature underneath the comment section to reveal the automated translation result, identifying the translation error, and categorizing them based on translation error category. The sentences written in captions were classified and analyzed, followed by interpreting the data and describing existing errors in the translation. To analyze the data, this study employed error categorization theory proposed by Koponen (2010) which led to the forming of reclassification of the relevant cases, namely:

1. Omitted concept: A certain word or sentence in ST is not described clearly in the TT text.
2. Added concept: ST text provides additional words or sentences when rendered into TT.
3. Untranslated concept: ST text has the same words in TT.
4. Mistranslated concept: ST text is translated incorrectly into TT.
5. Substituted concept: TT text represents an equivalent translation as a replacement due to the context.
6. Explicated concept: TT text explains direct or detailed information of ST.

Findings and Discussion

The translation tool provided by Instagram helps general users understand the captions in different languages. It can convert the source language of the caption into the language that is installed on the mobile device. When people open the application, the translation tool is already set automatically according to the general setting of the device. This research identified four types of machine translation errors on the captions of AjikDewa_43 Instagram account, namely omitted concept, untranslated concept, mistranslated concept, and substituted concept. The result of the analysis is shown in diagram 1 and a detailed discussion follows.

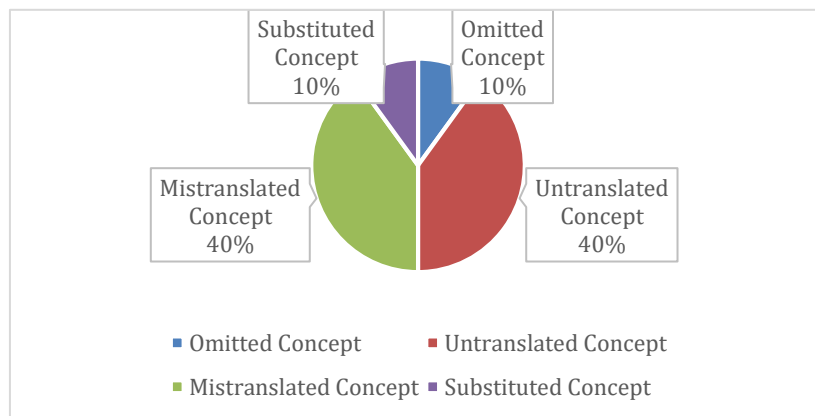


Diagram 1. Translation Error Analysis Result of AjikDewa_43 Instagram Account

1. Omitted Concept

Translation in omission usually pinpoints a missing source of words or a text. Examples of the mentioned concept can be viewed in table 1 below:

Table 1. Translation of Data 1

Source Text (ST)	Target Text (TT)
<p><i>“Penarikan pusaka di tempat angker dan dapat sandal dari @swallowfootwear colab @officialboncabe” Wah ni yang bagus dan cocok buat explore mentemen. Yuk buruan pesen.”</i></p>	<p>“Haunted trails and get sandals from @officialboncabe colab @swallowfootwear Wow this is good and suitable for explore mentemen. Let’s order now.”</p>

The translation in data 1 states an endorsement of flip flops by using horror theme content. The problem of the translation can be identified from the missing part of the phrase “*Penarikan pusaka di tempat angker*”. The machine translated it into “Haunted trails” only without translating “*penarikan pusaka*” into English. To correct the translation, the caption should be written as: “Inheritance withdrawal in a horror place and get flip flops from @officialboncabe in collaboration with @swallowfootwear. It is good and suitable for exploring, guys. Order it now.

2. Untranslated Concept

This process usually finds an untranslated word or sentence from the source language used in the target language. There are four data that can be elaborated according to the concept, which can be found in data 2 to 6.

Table 2. Translation of Data 2

Source Text (ST)	Target Text (TT)
<p><i>“Liu ade berita viral di Bali ternyata e. Patuh viralne care @cornerhouse_vape Di sana menerima jasa reparasi/service vape ya ton. Termurah dan terlengkap di Badung. Pokokne de ragu jak @cornerhouse_vape @cornerhouse_vape”</i></p>	<p>“Liu ade viral news in Bali apparently e. Follow the viral care @cornerhouse_vape There accepting repair/vape service ya ton. The cheapest and most complete in Badung. Of course they are raguing @cornerhouse_vape @cornerhouse_vape”</p>

Data 2 describes Balinese as the source language, translated to English as the target language. However, many Balinese terms from the source language are not translated and remain the same in the target language due to the fact that Instagram cannot convert a local language, such as Balinese, into English. This happens because Instagram automatically reads the language setting in the user’s device and the media does not provide a particular local language as one of the language options. Therefore, the translation resulted in an untranslated concept of texts or words as follows:

- a) “*Liu ade,*” in Balinese language means “*ada banyak*” in Indonesian language. Therefore, the correct English translation should be “There are many”.
- b) “*Ternyata e,*” the alphabet “e” in Balinese language is mostly used as an agreeing expression. The correct translation is erasing the “e” alphabet and keep the “apparently” word in the target text.
- c) “*Patuh viralne care,*” means “*sama seperti viralnya*”. It’s a complex Balinese sentence that should be written in English as “as viral as...”.

- d) “*Pokokne de ragu jak,*” means “*Pokoknya jangan ragu*”. The best English translation of this sentence is “At any rate, do not hesitate.”
- e) “*Ya ton*” is an additional expression used to ensure people that means “*ya, guys*”. The English translation is best written only with “*guys*” and erase the “*ya*”.

Based on the findings of the untranslated concept, the translation of the caption should be: “There are many viral news in Bali apparently. It’s as viral as the @cornerhouse_vape. It accepts vape service for you guys, the cheapest and most complete in Badung city. At any rate, do not hesitate to visit @cornerhouse_vape @cornerhouse_vape”.

Table 3. Translation of Data 3

Source Text (ST)	Target Text (TT)
“Jaman corona, sekarang punya usaha, sebaiknya punya QRIS BRI buat alat pembayaran. Tanpa bersentuhan langsung dengan nasabah. Syarat mudah: FOTO KTP, FOTO BUKU TABUNGAN BRI, DAN FOTO OLINE SHOP IG, NO WA. ”	“In the corona era, now having a business, it’s better to have QRIS BRI as payment tool. Without direct contact with the customer. The terms are simple: FOTO KTP, PICTURES OF BRI SAVINGS BOOK AND PHOTOS ONLINE SHOP IG, NO WA. ”

The caption in data 3 shows information about having a payment method using a barcode scanner in order to shorten direct contact with customers, especially in the pandemic era. The information from the source language is correctly written, but when it is translated into the target language, there are untranslated terms such as “FOTO KTP”, “QRIS BRI”, and “NO WA”. This is caused by the machine translation’s inability to generate the equivalent terms of the source language contextually and grammatically. Therefore, the best translation for the caption is: “In the pandemic era, QRIS BRI is mandatory in running a business as an alternative payment tool to reduce direct contact with customers. The terms are simple: ID CARD, PHOTO OF BRI SAVING ACCOUNT BOOK, PHOTO OF ONLINE SHOP ON IG, WHATSAPP NUMBER”.

Table 4. Translation of Data 4

Source Text (ST)	Target Text (TT)
“ <i>Ternyata lemah jantung cang nok. Buat semeton, ajik bakal live setiap hari di halaman fb kayak dulu lg ya. Biar kembali lagi halaman fb seperti sediakala. Nb: males main ig, baperan, dikit dikit main hapus.</i> ”	“Evidently my heart is weak. For semeton , ajik will live every day on fb page like before again ya. Let’s come back to the fb page as it is. Nb: lazy to play IG, baperan , will delete it in a bit.”

Translation in data 4 uses an informal style of communication according to the caption. Unfortunately, it shows several untranslated concepts in rendering the source language into the target language. For example, the source terms “*semeton*” and “*baperan*” are not translated into the target language. It remained the same in the target text. The word “*semeton*” is a Balinese word that means “relatives” while “*baperan*” is an Indonesian word that means “too sensitive”. To correct the whole translation accordingly, the caption should be written as: “It turned out my heart is really weak, man. For you guys, Ajik will be live every day on the Facebook page as before to reestablish the page like it was. NB: Too lazy to play IG, too sensitive, and sometimes likes to erase posts.”

Table 5. Translation of Data 5

Source Text (ST)	Target Text (TT)
<p>“<i>Tempat nongkrong murah meriah di jalan raya Ubud ne woi di @simplysocial.bali. Buat yg gak punya uang dan pengen banyak gaya, rekomen baged ni tempatnya. Jeg murah meriah.</i>”</p>	<p>“Cheap hangout spot on Ubud highway Di @simplysocial.bali. For those who don’t have money and want to have a lot of style, this is the place to recommend. Jeg murah meriah.”</p>

Data 5 shows another example of an untranslated concept. The expression “*jeg murah meriah*” means “it is cheap at any rate”. The Balinese word “*jeg*” is used to emphasize the meaning. However, it is not translated into the target language since Instagram machine translation does not recognize the local term. Therefore, the term remains untranslated in the target text. The suggested translation of the source caption is “A cheap hangout spot on Ubud Highway in @simplysocial.bali. For those who don’t have a lot of money and want to be a little showy. This place is very recommended for you who seek cheap menu”.

3. Mistranslated Concept

The translation process unsuccessfully renders the source language into the target language. Examples of mistranslated concepts can be seen in data 6 to data 9.

Table 6. Translation of Data 6

Source Text (ST)	Target Text (TT)
<p>“<i>Toleransi itu penting, jangan bicara menghargai kalau anda sendiri tidak bisa menghargai kami, kalian cuma oknum yang mencoreng nama orang orang Jawa/Muslim yang tinggal di Bali akhirnya jadi rusak cuma gara gara salah satu oknum seperti kalian. Kita sudah rukun di Bali. Tolong jangan dirusak. Salam toleransi.</i>”</p>	<p>“Tolerance is important. Don’t talk about respecting if you yourself can’t respect us, you are just oknum who spread the name of Javanian/Muslim who live in Bali finally become broken just because of one oknum like you. We already gathered in Bali. Please do not break it. Hello tolerance.”</p>

In data 6, the translated text shows several errors in rendering the source language into the target language even though the source caption uses Indonesian as the general language. An example of mistranslated concept can be seen in the translation of the term “*Orang Jawa*” into “Javanian” instead of “Javanese”. Furthermore, the source terms “*rukun*” and “*salam toleransi*” are translated into “gathered” and “Hello tolerance”, which are grammatically and semantically inaccurate. The translation for the two terms should be “live in harmony” and “Keep our tolerance”. The complete suggested translation for the source caption in data 6 is “Tolerance is important, do not talk about self-respect if you can’t respect us. You are just an individual who gives Javanese/Muslim citizens of Bali a bad name, which ended up being defective because of an individual like you. We already live in harmony here and don’t destroy it. Keep our tolerance.”

Table 7. Translation of Data 7

Source Text (ST)	Target Text (TT)
<p>“<i>Apa cuma ajik yang gak tau siapa sebenarnya mereka ini? Padahal dulu waktu masih SD suka sekali denger lagunya. Jalan melali ke hotel Nusa Dua. Kalau udah besar kira kira mereka masih melali ke hotel Nusa Dua gak?</i>”</p>	<p>“Is it only dogs who don’t know who these people are? Even though in the past when I was still <i>SD</i> I really liked to listen to the song. Walking to Nusa Dua Hotel. If they have grown up, do you think they still forget to go to hotel Nusa Dua?”</p>

In data 7, an example of a mistranslated concept is spotted in the translation of the word “*Ajik*” into “dogs” in the quote: “*Apa cuma ajik yang gak tau...*” to “Is it only dogs who don’t know...”. The word “*Ajik*” is a Balinese address term for fathers descended from the *Brahmin caste* (*Ida Bagus*) or *Ksatrya* (*Cokorda, Anak Agung, Gusti*). However, the Instagram machine translation mistranslates it into “dogs,” which is highly irrelevant to the context. Another example of a mistranslated concept can be seen in the translation of the Balinese word “*melali*” into “forget”. The word “*melali*” means “strolling around”. Therefore, translating it into “forget” can be categorized as a mistranslated concept. The suggested translation for this caption is: “Is it only me who really doesn’t know they are? I used to listen to their songs when I was still an elementary student. They stroll around Nusa Dua Hotel. When they have grown up, do you think they still play in the hotel?”

Table 8. Translation of Data 8

Source Text (ST)	Target Text (TT)
<p>“<i>Yuk ton daerah Penarungan dan lain sebagainya. Gaskan laundry di @ars_kopsafurmlaundry di sana laundry murah, terpercaya, cepat malas cang debat. Pokokne gaskan. Bisa paket expres 3 jam jadi.</i>”</p>	<p>“Yuk ton fighting area and so on. Apply laundry at @ars_kosparfumlaundry the laundry there is cheap, trusted, too lazy to debate. Anyway, just give it up. Can express package for 3 hours.”</p>

Data 8 shows a mistranslated concept from the source language into the target language. The main focus of this particular translation is “*daerah Penarungan*”. The sentence stated a specific location in one of Bali territories, but the machine translation incorrectly translated it into “fighting area”. The correct translation should be “Penarungan area”. To conclude the whole translation, the clearer rendering of translation is: “For those who live in Penarungan area and its surroundings, come visit laundry in @ars_kopsafurmlaundry. It is cheap, trusted, and I don’t have to tell you more. Just visit the place in any way, and you can order a 3-hour laundry express.”

Table 9. Translation of Data 9

Source Text (ST)	Target Text (TT)
“ <i>Sinchan nyari bakalan bonsai. Daripada kalian nyari mending beli di @nanoebonsaitaman Segala jenis bonsai ade ditu.</i> ”	“Sinchan is looking for a bonsai garden . Instead of you looking for it better buy it in @nanoebonsaitaman. All kinds of bonsai are there.”

The caption in data 9 shows an endorsement of one of the florists in Bali. The main focus is on the phrase “*bakalan bonsai*,” which is translated into “bonsai garden”. The phrase “*bakalan bonsai*” can be translated as “potential bonsai” or “bonsai material” in English. This refers to plants that have the potential to be made into bonsai after undergoing a special care and pruning process. Therefore, translating “*bakalan bonsai*” into “bonsai garden” is inaccurate and can be categorized as a mistranslated concept. In order to evaluate the translation, the best result in rewriting the caption is: “Sinchan is searching for a bonsai plant. Instead of looking elsewhere, you better buy one in @nanoebonsaitaman. It sells all kinds of bonsai plants.

4. Substituted Concept

The concept or the meaning of the translated word is changed but ineligible. An example of a substituted concept can be seen in table 10 below:

Table 10. Translation of Data 10

Source Text (ST)	Target Text (TT)
“ <i>Alasan 17 pintu masuk Pantai Kuta ditutup.</i> ”	“The reason 17 entrance to the city beach is closed.”

The short caption as shown in data10 shows a humoric context of the content, but the translation result is not accurate even though the written text is short. The phrase “*Pantai Kuta*” is substituted by “the city beach”. It shows an ineligible word replacement from the source language because “*Pantai Kuta*” should be translated into “Kuta Beach”. The suggested translation for the caption is “The reason why the 17 entrances to Kuta Beach are closed.”

Conclusion

The result of the analysis shows that there are several errors found in the translation result of the Instagram translation machine in rendering Ajikdewa_43 captions. The errors found were omitted concept, untranslated concept, mistranslated concept, and substituted concept. Problems in Instagram translation machine occur because of the random analysis, it causes many mistranslations of the contextual meaning of the texts or the captions. The data analysis shows how Instagram translation machine fails to translate Indonesian and Balinese terms such as “*semeton*”, “*melali*”, and “*baperan*”. Those common words can be translated correctly if the machine translation facilitates more local languages. Therefore, it can be concluded that the translation feature on Instagram captions is considered an additional facility to help users understand foreign languages. However, the machine still needs improvements in order to produce more accurate translations.

References

- Adawiyah, A. R., Baharuddin, B., Wardana, L. A., & Farmasari, S. (2023). Comparing Post-Editing Translations by Google NMT and Yandex NMT. *Teknosastik: Jurnal Bahasa dan Sastra*, 21(1), 23-34.
- Anisah, N., Sartika, M., & Kurniawan, H. (2021). Penggunaan Media Sosial Instagram dalam Meningkatkan Literasi Kesehatan Pada Mahasiswa. *Jurnal Peurawi: Media Kajian Komunikasi Islam*, 4(2), 94-107.
- Antika, R. (2022). Translation Shifts and Equivalence Strategy Produced By Instagram Machine Translation. *JIPIS*, 31(1), 63-73.
- Astuti, S. R. T., Khasanah, I., & Yoestini, Y. (2020). Study of impulse buying on Instagram users in Indonesia. *Diponegoro International Journal of Business*, 3(1), 47-54.
- Darmaningrat, E. W. T., Astuti, H. M., & Alfi, F. (2020). Information Privacy Concerns Among Instagram Users: The Case of Indonesian College Students. *Journal of Information Systems Engineering and Business Intelligence*, 6(2), 159-168.
- Faruk. (2012). *Metode Penelitian Sastra*. Yogyakarta: Pustaka Pelajar.
- Kendenan, E. S., Kristen, U., & Wacana, S. (2023). An Analysis of Politeness Strategy Used by The Actors of Johnny English Film and Its Translation. *Teknosastik: Jurnal Bahasa dan Sastra*, 21(2), 102-109.
- Koponen, M. (2010). Assessing Machine Translation Quality with Error Analysis. *Electronic Proceedings of the VIII KäTu Symposium on Translation and Interpreting Studies*, 4, 1-12.
- Leung, F. F., Gu, F. F., Li, Y., Zhang, J. Z., & Palmatier, R. W. (2022). Influencer Marketing Effectiveness. *Journal of Marketing*, 86(6), 93-115.
- Meilasari, P. (2019). Instagram Translation Machine: How Problematic Is It? *Widya Warta*, 1(1), 124-133.
- Nadilla, R., & Lubis, S. (2023). Lexical Errors Made by Instagram Machine Translation in Translating the Account of “CNN Indonesia” News Article. *LingPoet: Journal of Linguistics and Literary Research*, 4(1), 30-45.
- Natasha, J., Angelia, C. R., & Susilo, D. (2023). Social Media Marketing Strategy in Increasing Customer Engagement of Taman Safari Indonesia Bogor (Case Study on Instagram

@Taman_Safari). *Indonesian Journal of Business Analytics*, 3(2), 351-408.

- Putri, A. T., & Setiajid, H. H. (2021). Instagram Translate and Human Translation in the English Captions of Jokowi's Account: An Analysis of Koponen's Error Category. *4th English Language and Literature International Conference (ELLiC) Proceedings*, 4, 432-436.
- Rizaty, M. (2022). *Pengguna Instagram Indonesia Terbesar Keempat Di Dunia*. DataIndonesia.Id. <https://dataindonesia.id/digital/detail/pengguna-instagram-indonesia-terbesar-keempat-di-dunia>
- Swadeshi, A. W., & Sutrisno, A. (2021). Lexical Error Analysis of Indonesian-English Translation of Texts in Dewantara Kirti Griya Museum. *Lexicon*, 8(2), 105-114.
- Ting, H., Ming, W. W. P., De Run, E. C., & Choo, S. L. Y. (2015). Beliefs about the Use of Instagram: An Exploratory Study. *International Journal of Business and Innovation*, 2(2), 15-31.