Journal of Information Systems and Informatics

**Vol. 6, No. 4, December 2024** e-ISSN: **2656-4882** p-ISSN: **2656-5935**

DOI: 10.51519/journalisi.v6i4.918 Published By DRPM-UBD

# Comparison of Conversational Corpus and News Corpus on Gender Bias in Indonesian-English Transformer Model Translation

**Andik Wijanarko[1], Adzkiyatun Nisa Al Haura[2], Indar Puspitaningrum[3], Dhanar Intan Surya Saputra[4]**

[1]Information Technology Departement , Amikom Purwokerto University, Banyumas, Indonesia
[2,3,4]Informatics Departement, Amikom Purwokerto University, Banyumas, Indonesia
Email: [1]andikwijanarko@amikompurwokerto.ac.id, [2]adzkiyaanis4@gmail.com,
[3]indar.puspitaningrum16@gmail.com, [4]dhanarsaputra@amikompurwokerto.ac.id

## Abstract

Gender bias in machine translation is a significant issue that affects text translation and gender perception, often leading to misunderstandings, such as the tendency to default to using male pronouns. For example, the word "*dia*" in Indonesian is often translated as "he" rather than "she," even when the context suggests otherwise, as seen in the case of President Megawati. Reducing this bias requires ongoing research, particularly in understanding how different corpora affect translation accuracy. Studies have shown that formal news corpora, which have less gender bias, produce different results compared to conversational corpora that are more informal and exhibit gender bias. This research uses a training dataset of the Indonesian-English conversational parallel corpus from Open Subtitles, which contains many gendered pronouns. Additionally, a news corpus from Tanzil, with fewer gendered words, was also used. These corpora were sourced from Opus, widely used by previous researchers. For the testing dataset, biographies of female presidents were used, which are often translated as masculine by popular machine translation systems by default. Each corpus was trained using a Transformer model, resulting in a translation model. Each sentence from the generated translations was then detected for gender, and compared with the gender of sentences from the test data to evaluate accuracy. The results showed that the accuracy of gender translation from the conversational corpus was 84%, while the news corpus achieved an accuracy of 8%.

**Keywords**: Machine Translation, Gender Bias, Transformer, Indonesia-English.

## 1. INTRODUCTION

Gender bias is a specific issue in machine translation, for example, Google's translation engine, which has been criticized for its default masculine gender bias. The default masculine bias in Google Translate is predominant in the context of STEM (Science, Technology, Engineering, and Mathematics) professions, at 72% [1].

This creates gender stereotypes related to the spread of negative social group generalizations, such as devaluing feminine representation in less prestigious jobs, like associating "teacher" with feminine and "lecturer" with masculine [2]. Such behavior can be harmful as it affects the self-esteem of the target group [3]. If women are the subject of a traditionally masculine job in text automatically translated by a machine translator, they are likely to be identified as men, leading readers to experience gender inaccuracies [4]. Moreover, users may not notice gender errors in machine translation due to the system's fluent output[5] which, over time, can reinforce stereotypical assumptions and biases (e.g., only men being qualified for high-level positions) [6].

The phenomenon of gender bias in machine translation can be assessed by mapping sentences constructed in gender-neutral languages, such as Malay, Estonian, Finnish, Yoruba, and others, to English using automatic translation tools. Indonesian is a gender-neutral language, similar to Malay. The Indonesian-English translation results from Google Translate tend to identify the word *'dia'* in Indonesian as male, translating it as 'he' in English rather than as female, translating it as 'she.' Even when the previous sentence clause provides context for 'he/she,' the translation still tends to default to 'he' [7]. Another example of gender-biased translation error in Indonesian-English on Google Translate is the sentence, "*Presiden Megawati, dia adalah presiden ke-5 Indonesia*." The word "*dia*," which should be translated as "she," is instead translated as "he." This is a type of stereotypical error where the pronoun for "President" defaults to "he."

Machine translation refers to software that can automatically convert source sentences into target sentences. It is highly beneficial for translating between languages, such as from a foreign language to Indonesian or vice versa [8]. However, both human and machine translation face various challenges, particularly when translations need to be contextually accurate. These challenges are influenced by factors like geographic location, culture, habits, and lexical differences[9]. Machine translation methods have evolved significantly, starting from Rule-Based Machine Translation (RBMT)[10], [11], [12], progressing to Statistical Machine Translation (SMT) [13], [14], [15], and now advancing to the state-of-the-art Neural Machine Translation (NMT), which uses artificial neural networks [16], [17], [18]

The NMT approach is not limited to performing language translation; it also encompasses a model structure and a process layer comprising the methods used. NMT utilizes a model called Sequence to Sequence (SeqToSeq) [19], which facilitates the language translation process. The SeqToSeq model operates in two stages: Encoder and Decoder [20]. The Encoder is a process layer that inputs the source language, while the Decoder is a process layer that translates the output from the Encoder into the target language or translated language [21]. The

Encoder-Decoder layer employs a learning process network based on the Recurrent Neural Network (RNN) method [22].

RNN is a commonly used method for processing sequential data, such as text processing and other similar tasks [23]. An RNN in NMT consists of three layers: an input layer, which maps each word to a vector using word embeddings or a one-hot word index; a recurrent hidden layer, which continuously calculates and updates the hidden state after reading each word; and an output layer, which estimates the probability of the next words based on the current hidden state [24]. The RNN scheme is illustrated in Figure 1.

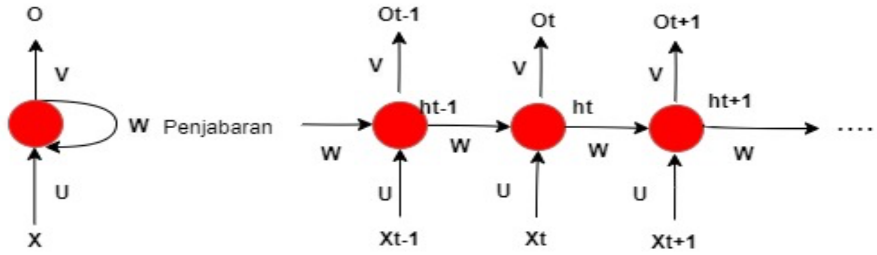


**Figure 1.** RNN Scheme [24].

Encoder-Decoder is a framework for the Sequence-to-Sequence method, comprising two main components. The first part is the Encoder, which vectorizes words from the input sentence, and the second part is the Decoder, which predicts the translated words based on the vectorized values obtained from the Encoder [25]. The Encoder-Decoder scheme is illustrated in Figure 2.

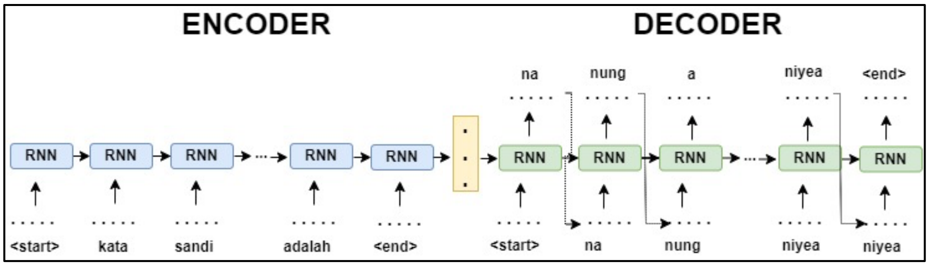


**Figure 2.** Encoder-Decoder Scheme

Each cell of the RNN model in the Encoder-Decoder schematic image in Figure 3 above uses the (Gated Recurrent Unit) model. Most competitive sequence-to-sequence neural models have an Encoder-Decoder structure. Here, the Encoder maps the input sequence of symbols $(x_1, ..., x_n)$ to a continuous representation sequence $z = (z_1, ..., z_n)$.. Given z, the Decoder then generates the output sequence $(y_1, ..., y_n)$ of symbols, one element at a time. At each step, the model is autoregressive, using previously generated symbols as additional input when producing the next step. The Transformer follows this overall architecture by using stacked self-attention and point-wise, fully connected layers for both the Encoder and Decoder. The Transformer schema can be seen in Figure 3.

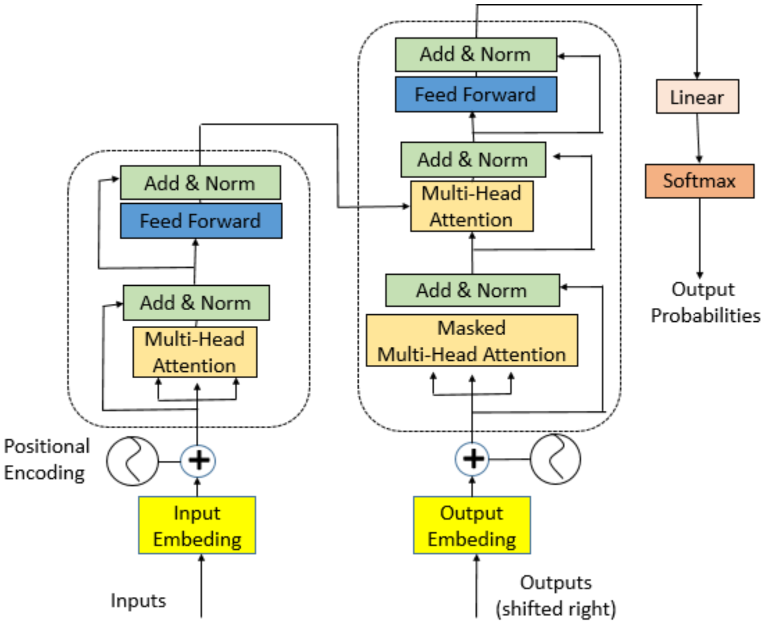


**Figure 3.** The Transformer Model

## 2. METHODS

This research was carried out experimentally, starting from collecting data in the form of a parallel corpus of Indonesian-English languages, then pre-processing the corpus and then training using the Transformer model to produce a translation model. The translation model is used to translate Indonesian sentences into English. The final step is to test the translation results using a model with translation results carried out by humans (human translators) using the BLEU measurement standard. The research procedure can be seen in the following picture.

The process involves several steps, starting with downloading the conversational and news corpora from the OPUSNLP website. Next, each corpus is divided into three datasets: 60% for training, 20% for validation, and 20% for testing. The datasets are then trained using the Transformer model. After training, translation is performed on the test data using their respective models. The translations produced by the models are evaluated by comparing them with the original translations in the corpus, using the BLEU method for assessment. Finally, a gender accuracy test is conducted on the test data, and the results are compared with the gender accuracy from the corpus test data using accuracy metrics.

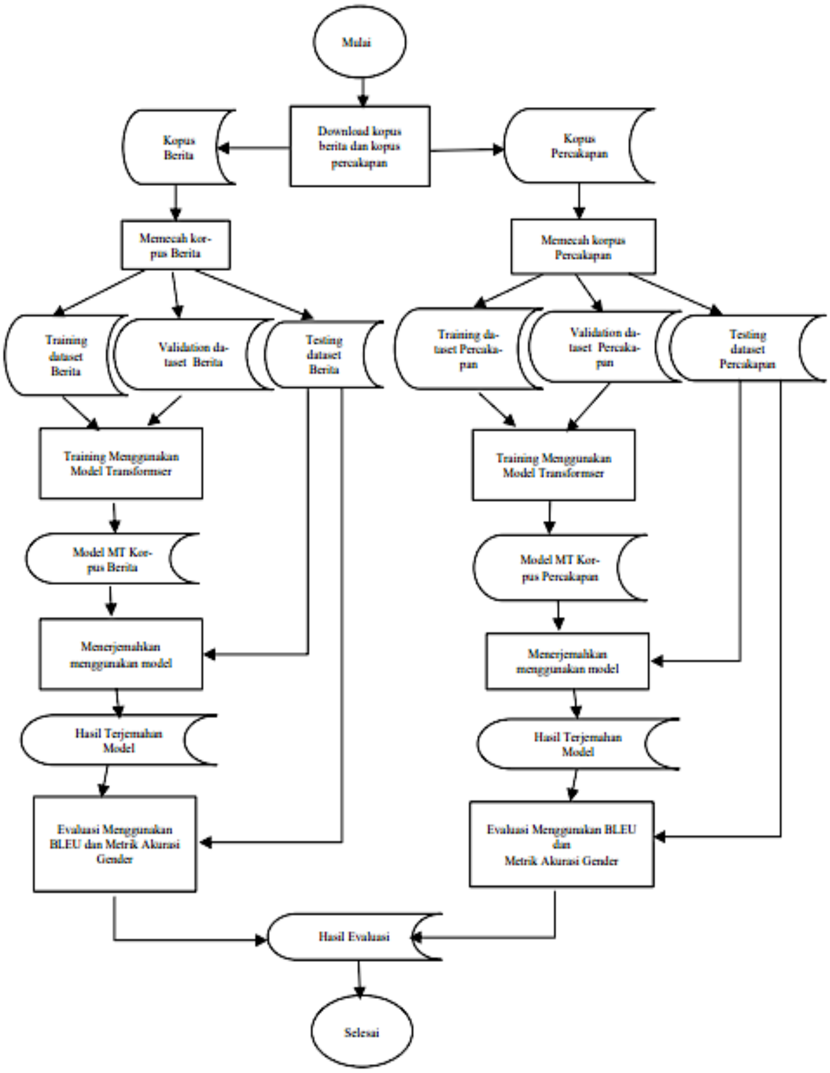

Mulai
Download kopus berita dan kopus percakapan
Kopus Berita
Kopus Percakapan
Memecah kor-pus Berita
Memecah korpus Percakapan
Training dataset Berita
Validation da-taset Berita
Testing dataset Berita
Training da-taset Percaka-pan
Validation da-taset Percaka-pan
Testing dataset Percakapan
Training Menggunakan Model Transformser
Training Menggunakan Model Transformser
Model MT Kor-pus Berita
Model MT Kor-pus Percakapan
Menerjemahkan menggunakan model
Menerjemahkan menggunakan model
Hasil Terjemahan Model
Hasil Terjemahan Model
Evaluasi Menggunakan BLEU dan Metrik Akurasi Gender
Evaluasi Menggunakan BLEU dan Metrik Akurasi Gender
Hasil Evaluasi
Selesai

**Figure 5.** Research Procedure

## 2.1. Create Parallel Corpus

The Indonesian-English language parallel corpus was created by translating Indonesian sentences into English language. The parallel corpus contains 100,000 pairs of Indonesian-English sentences.

### 2.2. Pre-processing

Before training is carried out on the dataset using the model, pre-processing is carried out on the parallel corpus with the aim of making the training process more optimal. The types of pre-processing carried out in this research include: Lowercase, Lowercase changes all characters to non-capital letters, this aims to ensure that the same word is truly the same and is not differentiated by letters. For example, the words good and good are the same word. Lowercase is a pre-processing standard for machine translation which results in a smaller number of unique words, thereby reducing features. Punctuation, this process removes punctuation marks in all sentences. The aim of removing punctuation marks is so that there are no additional words that contain punctuation marks so that the number of unique words that are features becomes smaller. This will result in the training process being faster and more accurate. Just like lowercase, punctuation is widely used as a pre-process in building machine translation.

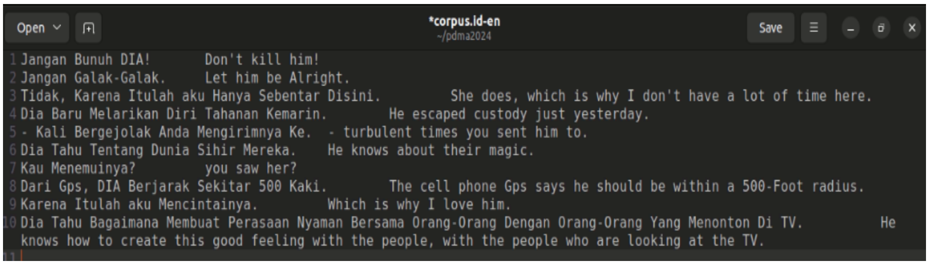
### 2.3. Training Corpus using Transformer

The main step of this research is the training dataset. The dataset in the form of a parallel corpus is trained using the Transformer model. The Transformer model has been used several times to build machine translations such as German-English machine translations [20], French, Arabic and Chinese to Urdu, Chinese-English with varying results.

## 3. RESULTS AND DISCUSSION

### 3.1. Dataset Description

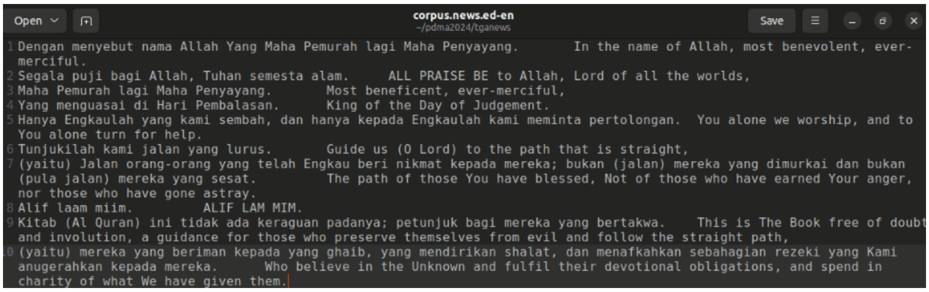
The dataset used is in the form of a parallel corpus, namely a pair of Indonesian and English languages where on average each Indonesian sentence contains 10 words. Each Indonesian sentence is paired with a English sentence with a separating sign in the form of Tab. Parallel corpus artifact can be seen in Figure 4.



```
Jangan Bunuh DIA!       Don't kill him!
Jangan Galak-Galak.     Let him be Alright.
Tidak, Karena Itulah aku Hanya Sebentar Disini.         She does, which is why I don't have a lot of time here.
Dia Baru Melarikan Diri Tahanan Kemarin.        He escaped custody just yesterday.
- Kali Bergejolak Anda Mengirimnya Ke.  - turbulent times you sent him to.
Dia Tahu Tentang Dunia Sihir Mereka.    He knows about their magic.
Kau Menemuinya?         you saw her?
Dari Gps, DIA Berjarak Sekitar 500 Kaki.        The cell phone Gps says he should be within a 500-Foot radius.
Karena Itulah aku Mencintainya.         Which is why I love him.
Dia Tahu Bagaimana Membuat Perasaan Nyaman Bersama Orang-Orang Dengan Orang-Orang Yang Menonton Di TV.          He
knows how to create this good feeling with the people, with the people who are looking at the TV.
```

**Figure 4.** Conversational Parallel Corpus

Figure 4 above is a parallel Indonesian-English conversational corpus. The parallel corpus contains 100,000 pairs of Indonesian English sentences with an average length of Indonesian sentences of 10 words. In addition, a news corpus was used for comparison, as shown in Figure 5. Figure 6 shows that similar to the conversation corpus, the news corpus consists of 100,000 Indonesian-English sentence pairs.



```
Dengan menyebut nama Allah Yang Maha Pemurah lagi Maha Penyayang.       In the name of Allah, most benevolent, ever-
merciful.
Segala puji bagi Allah, Tuhan semesta alam.     ALL PRAISE BE to Allah, Lord of all the worlds,
Maha Pemurah lagi Maha Penyayang.       Most beneficent, ever-merciful,
Yang menguasai di Hari Pembalasan.      King of the Day of Judgement.
Hanya Engkaulah yang kami sembah, dan hanya kepada Engkaulah kami meminta pertolongan.  You alone we worship, and to
You alone turn for help.
Tunjukilah kami jalan yang lurus.       Guide us (O Lord) to the path that is straight,
(yaitu) Jalan orang-orang yang telah Engkau beri nikmat kepada mereka; bukan (jalan) mereka yang dimurkai dan bukan
(pula jalan) mereka yang sesat.         The path of those You have blessed, Not of those who have earned Your anger,
nor those who have gone astray.
Alif laam miim.         ALIF LAM MIM.
Kitab (Al Quran) ini tidak ada keraguan padanya; petunjuk bagi mereka yang bertakwa.    This is The Book free of doubt
and involution, a guidance for those who preserve themselves from evil and follow the straight path,
(yaitu) mereka yang beriman kepada yang ghaib, yang mendirikan shalat, dan menafkahkan sebahagian rezeki yang Kami
anugerahkan kepada mereka.      Who believe in the Unknown and fulfil their devotional obligations, and spend in
charity of what We have given them.
```

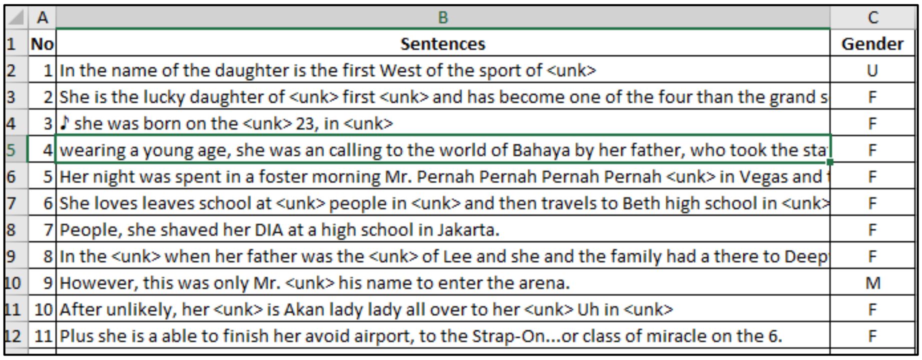
**Figure 5.** News Parallel Corpus

### 3.2. TrainingDatasets

Parallel corpus training uses hardware devices with Intel Core 15 Gen 11 processor specifications, 16 GB RAM, 500 GB SSD. The operating system software is Linux Ubuntu 22, Python 3 programming language. Dataset training time is 7 days using a GPU (Graphic Processing Unit).

### 3.3. Evaluation

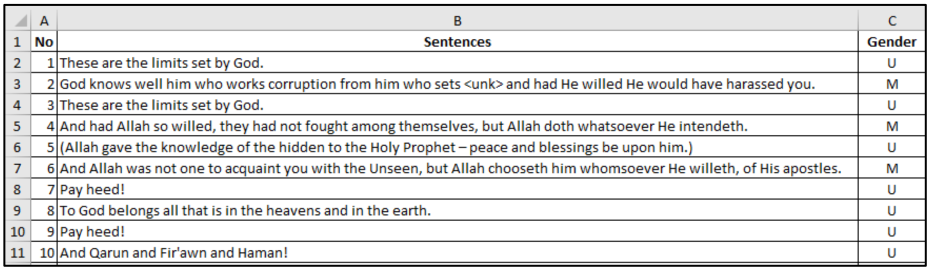
The first step in the evaluation phase is to identify and annotate the gender of the sentences produced by the model. Sentences containing the gendered word 'She' will be identified as Female (F), while those containing the word 'He' will be identified as Male (M), as shown in Figure 6.



| | A | B | C |
|---|---|---|---|
| 1 | No | Sentences | Gender |
| 2 | 1 | In the name of the daughter is the first West of the sport of <unk> | U |
| 3 | 2 | She is the lucky daughter of <unk> first <unk> and has become one of the four than the grand s | F |
| 4 | 3 | ♪ she was born on the <unk> 23, in <unk> | F |
| 5 | 4 | wearing a young age, she was an calling to the world of Bahaya by her father, who took the sta | F |
| 6 | 5 | Her night was spent in a foster morning Mr. Pernah Pernah Pernah Pernah <unk> in Vegas and | F |
| 7 | 6 | She loves leaves school at <unk> people in <unk> and then travels to Beth high school in <unk> | F |
| 8 | 7 | People, she shaved her DIA at a high school in Jakarta. | F |
| 9 | 8 | In the <unk> when her father was the <unk> of Lee and she and the family had a there to Deep | F |
| 10 | 9 | However, this was only Mr. <unk> his name to enter the arena. | M |
| 11 | 10 | After unlikely, her <unk> is Akan lady lady all over to her <unk> Uh in <unk> | F |
| 12 | 11 | Plus she is a able to finish her avoid airport, to the Strap-On...or class of miracle on the 6. | F |

**Figure 6.** Gender annotation of the model's conversation translation

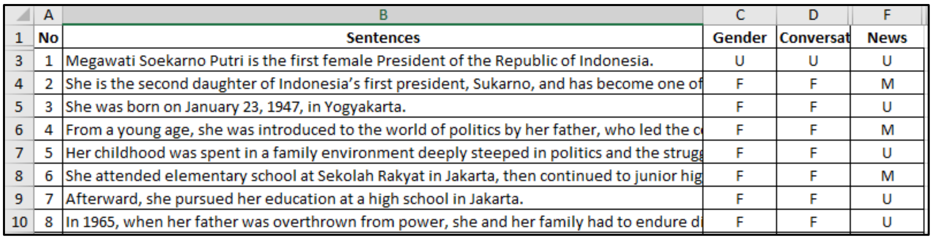
The Figure 6 shows the gender annotation of sentences translated by the model from the training corpus of conversations.



| | A | B | C |
|---|---|---|---|
| 1 | No | Sentences | Gender |
| 2 | 1 | These are the limits set by God. | U |
| 3 | 2 | God knows well him who works corruption from him who sets <unk> and had He willed He would have harassed you. | M |
| 4 | 3 | These are the limits set by God. | U |
| 5 | 4 | And had Allah so willed, they had not fought among themselves, but Allah doth whatsoever He intendeth. | M |
| 6 | 5 | (Allah gave the knowledge of the hidden to the Holy Prophet – peace and blessings be upon him.) | U |
| 7 | 6 | And Allah was not one to acquaint you with the Unseen, but Allah chooseth him whomsoever He willeth, of His apostles. | M |
| 8 | 7 | Pay heed! | U |
| 9 | 8 | To God belongs all that is in the heavens and in the earth. | U |
| 10 | 9 | Pay heed! | U |
| 11 | 10 | And Qarun and Fir'awn and Haman! | U |

**Figure 7**. Gender annotation of the model's news translation

The Figure 7 shows the gender annotation of sentences translated by the model from the training corpus of news. The gender annotation results from both model translations are then compared with the gender annotations in the test data to obtain the results, as shown in Figure 8.



| | A | B | C | D | F |
|---|---|---|---|---|---|
| 1 | No | Sentences | Gender | Conversat | News |
| 3 | 1 | Megawati Soekarno Putri is the first female President of the Republic of Indonesia. | U | U | U |
| 4 | 2 | She is the second daughter of Indonesia's first president, Sukarno, and has become one of | F | F | M |
| 5 | 3 | She was born on January 23, 1947, in Yogyakarta. | F | F | U |
| 6 | 4 | From a young age, she was introduced to the world of politics by her father, who led the c | F | F | M |
| 7 | 5 | Her childhood was spent in a family environment deeply steeped in politics and the strugg | F | F | U |
| 8 | 6 | She attended elementary school at Sekolah Rakyat in Jakarta, then continued to junior hig | F | F | M |
| 9 | 7 | Afterward, she pursued her education at a high school in Jakarta. | F | F | U |
| 10 | 8 | In 1965, when her father was overthrown from power, she and her family had to endure di | F | F | U |

**Figure 8**. Compare Gender annotation of the model's

After the calculation, the model translation trained on the conversation corpus shows a gender translation accuracy rate of 84 percent, while the model trained on the news corpus shows an accuracy rate of 8 percent.

## 3.4. Discussion

This study investigated the influence of training datasets on machine translation models, specifically examining the translation accuracy of gendered pronouns across two distinct corpus types: conversational and news. The findings illuminate several key aspects regarding the suitability of dataset types for specific translation tasks and raise critical questions about model training practices in natural language processing (NLP).

The stark contrast in gender translation accuracy between the conversational (84%) and news (8%) models suggests that the inherent characteristics of the dataset play a crucial role in model performance. The conversational corpus likely provided richer context cues related to gender, given its dynamic and expressive nature. Such contexts assist the model in learning and predicting gendered pronouns more accurately. On the other hand, the structured and formal nature of news writing, which often omits detailed personal narratives or explicit gender identifiers, might contribute to the poor performance in gender-specific translations observed in the news corpus model.

The methodology used for annotating gender by identifying specific gendered pronouns ("She" as Female and "He" as Male) highlights a simplistic approach that may not capture the full spectrum of gender expressions, especially in languages and texts where gender neutrality or fluidity is present. This raises questions about the adequacy of such binary gender annotations in training datasets, potentially leading to biased or inaccurate translation models that do not reflect the diversity of gender expressions.

The differential performance of translation models based on dataset types also has broader implications for AI fairness and the ethical use of machine translation in global communication. If translation models are trained on data that does not adequately represent diverse linguistic and cultural contexts, there is a risk of perpetuating biases and misrepresentations in automated translations. This study underscores the need for careful selection and preparation of training data that not only improves model performance but also aligns with ethical standards in AI applications.

Future research could explore more nuanced and inclusive methods of gender annotation, perhaps incorporating non-binary and gender-neutral pronouns to enhance the models' understanding and representation of gender. Additionally, extending this analysis to other linguistic features and datasets, including those from different languages and cultural contexts, could provide deeper insights into the universal applicability of these findings and help in developing more robust and culturally sensitive translation models. Overall, the points toward the necessity for more deliberate and ethical approaches in training data selection and model development in NLP, emphasizing that technological advancements should go hand in hand with considerations of inclusivity and fairness.

## 4. CONCLUSION

The Transformer model can be used to build a prototype machine translation for Indonesian-English. The model is also capable of translating gender-biased words with varying accuracy, depending on the type of corpus used for the dataset. The conversation corpus type demonstrates a higher accuracy rate in translating gender-biased words compared to the news corpus type.

## 5. ACKNOLEDGEMENT

This research was funded by Universitas Amikom Purwokerto Penelitian Dosen Muda Amikom (PDMA) Tahun Anggaran 2024.

## REFERENCES

[1] M. O. R. Prates, P. H. Avelar, and L. C. Lamb, "Assessing gender bias in machine translation: a case study with Google Translate," *Neural Comput Appl*, vol. 32, no. 10, pp. 6363–6381, 2020, doi: 10.1007/s00521-019-04144-6.

[2] B. Savoldi, M. Gaido, L. Bentivogli, M. Negri, and M. Turchi, "Gender bias in machine translation," *Trans Assoc Comput Linguist*, vol. 9, pp. 845–874, 2021, doi: 10.1162/tacl_a_00401.

[3] D. Bourguignon, V. Y. Yzerbyt, C. P. Teixeira, and G. Herman, "When does it hurt? Intergroup permeability moderates the link between discrimination and self-esteem.," *European Journal of Social Psychology, 45(1):3–9.*, 2015.

[4] L. Zimman, E. Hazenberg, and M. Meyerhoff, "Trans peoples linguistic self-determination and the dialogic nature of identity," in *Linguistic, legal and everyday perspectives, pages 226–248.*, 2017.

[5] M. J. Martindale and C. Park, "Fluency Over Adequacy : A Pilot Study in Measuring User Trust in Imperfect MT," in *Proceedings of the 13th Conference of the Association for Machine Translation in the Americas (Volume 1: Research Track), pages 13–25, Boston, USA. Association for Machine Translation in the Americas*, 2018.

[6] I. Régner, C. Thinus-blanc, A. Netter, T. Schmader, and P. Huguet, "Committees with implicit biases promote fewer women when they do not believe gender bias exists," *Nature Human Behaviour, 3(11):1171–1179.*, 2019, doi: 10.1038/s41562-019-0686-3.

[7] Nurtamin, H. Abbas, E. Iswary, and M. Hasyim, "Gender Bias In Machine Translation ( Google Translate ) From Indonesian To English," *Journal of Positive School Psychology*, vol. 6, no. 4, pp. 9754–9761, 2022.

[8] A. Faqih, "Penggunaan Google Translate Dalam Penerjemahan Teks Bahasa Arab Ke Dalam Bahasa Indonesia," *ALSUNIYAT: Jurnal Penelitian Bahasa, Sastra, dan Budaya Arab*, vol. 1, no. 2, pp. 88–97, 2018, doi: 10.17509/alsuniyat.v1i2.24216.

[9] J. Sheny *et al.*, "The source-target domain mismatch problem in machine translation," *ArXiv*, 2019.

[10] A. Alqudsi, N. Omar, and K. Shaker, "A Hybrid Rules and Statistical Method for Arabic to English Machine Translation," in *2nd International Conference on Computer Applications and Information Security, ICCAIS 2019*, IEEE, 2019. doi: 10.1109/CAIS.2019.8769545.

[11] M. Singh, R. Kumar, and I. Chana, "Improving Neural Machine Translation Using Rule-Based Machine Translation," *2019 7th International Conference on Smart Computing and Communications, ICSCC 2019*, pp. 1–5, 2019, doi: 10.1109/ICSCC.2019.8843685.

[12] J. Zhang, M. Utiyama, E. Sumita, G. Neubig, and S. Nakamura, "Improving neural machine translation through phrase-based soft forced decoding," *Machine Translation*, vol. 34, no. 1, pp. 21–39, 2020, doi: 10.1007/s10590-020-09244-y.

[13] L. Li, C. Parra Escartín, A. Way, and Q. Liu, "Combining translation memories and statistical machine translation using sparse features," *Machine Translation*, vol. 30, no. 3–4, pp. 183–202, 2016, doi: 10.1007/s10590-016-9187-6.

[14] P. Koehn, *Statistical Machine Translation*, no. 2. 2017. doi: 10.5565/rev/tradumatica.203.

[15] H. Cuong and K. Sima'an, *A survey of domain adaptation for statistical machine translation*, vol. 31, no. 4. Springer Netherlands, 2017. doi: 10.1007/s10590-018-9216-8.

[16] M. A. Haji Sismat, "Neural and Statistical Machine Translation: A comparative error analysis," in *Conference: 17th International Conference on Translation*, 2019.

[17] Y. Zhang and G. Liu, "Paragraph-Parallel based Neural Machine Translation Model with Hierarchical Attention," *J Phys Conf Ser*, vol. 1453, no. 1, 2020, doi: 10.1088/1742-6596/1453/1/012006.

[18] J. E. Ortega, R. Castro Mamani, and K. Cho, "Neural machine translation with a polysynthetic low resource language," *Machine Translation*, vol. 34, no. 4, pp. 325–346, 2021, doi: 10.1007/s10590-020-09255-9.

[19] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," *Adv Neural Inf Process Syst*, vol. 4, no. January, pp. 3104–3112, 2014.

[20] B. Van Merri and C. S. Fellow, "Learning Phrase Representations using RNN Encoder – Decoder for Statistical Machine Translation," in *Proceedings ofthe 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2014, pp. 1724–1734.

[21] S. Garg, S. Peitz, U. Nallasamy, and M. Paulik, "Jointly learning to align and translate with transformer models," *EMNLP-IJCNLP 2019 - 2019 Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing, Proceedings of the Conference*, pp. 4453–4462, 2019, doi: 10.18653/v1/d19-1453.

[22] D. Britz, A. Goldie, M. T. Luong, and Q. V. Le, "Massive exploration of neural machine translation architectures," *EMNLP 2017 - Conference on Empirical Methods in Natural Language Processing, Proceedings*, pp. 1442–1451, 2017, doi: 10.18653/v1/d17-1151.

[23] Y. Liu, D. Zhang, L. Du, Z. Gu, J. Qiu, and Q. Tan, "A simple but effective way to improve the performance of RNN-Based encoder in neural machine translation task," in *Proceedings - 2019 IEEE 4th International Conference on Data Science in Cyberspace, DSC 2019*, IEEE, 2019, pp. 416–421. doi: 10.1109/DSC.2019.00069.

[24] X. Wang, C. Chen, and Z. Xing, "Domain-specific machine translation with recurrent neural network for software localization," *Empir Softw Eng*, vol. 24, no. 6, pp. 3514–3545, 2019, doi: 10.1007/s10664-019-09702-z.

[25] L. Corallo, G. Li, K. Reagan, A. Saxena, A. S. Varde, and B. Wilde, "A Framework for German-English Machine Translation with GRU RNN," in *CEUR Workshop Proceedings*, 2022.