Citation: Ullah, Z., Alam, M. F., & Ihsan, W. (2022). Juxtaposing Urdu to English Captions of Facebook/ Metaverse: A Neural Machine Translation Study. *Global Language Review, VII*(II), 55-66. https://doi.org/10.31703/glr.2022(VII-II).05

- **DOI:** 10.31703/glr.2022(VII-II).05
- **URL:** http://dx.doi.org/10.31703/glr.2022(VII-II).05
- **Vol.** VII, **No.** II (Spring 2021)
- **Pages:** 55 – 66

Juxtaposing Urdu to English Captions of Facebook/ Metaverse: A Neural Machine Translation Study

Zafar Ullah
Instructor, Virtual University, Islamabad, Pakistan.
Email: zafarullah@vu.edu.pk
(Corresponding Author)

Muhammad Farooq Alam
Assistant Professor, University of NUML, Rawalpindi, Punjab, Pakistan.

Wishma Ihsan
MPhil English Scholar, University of NUML, Islamabad, Pakistan.

- **p-ISSN:** 2663-3299
- **e-ISSN:** 2663-3841
- **L-ISSN:** 2663-3299

Abstract: *Facebook/ Metaverse is a world-leading social networking platform that is equally popular in the world that is equally popular globally Due to the diversity of its users, Facebook is dedicated to improving its Machine Translation (MT) to refine the communication process and remove linguistic hurdles for efficient communication. Recently, Metaverse has started Neural Machine Translation (NMT) for the improvement of translation. NMT uses artificial neural networks to predict a more appropriate word choice and sequence. This research investigates the efficiency of Neural Machine Translation (NMT) in translating text from Urdu to the English language. The Oxford Urdu English Dictionary and FARHANG E AASFIA dictionaries have been used as validation parameters in translation. The semantic and syntactic errors have been segregated and categorized with Antoine Berman's Twelve Deforming Tendencies for rectification. Major findings reveal that the most frequent errors result from transliteration that damages the semantic and syntactic structure of the text.*

Key Words: Facebook/ Metaverse, Machine Translation (MT), Neural Machine Translation (NMT), Artificial Neural Networks

Introduction

Facebook/ Metaverse is one of the leading social networking sites having over 2.6 billion active users worldwide with 101 language translations available. One of Facebook's main aims is to remove obstacles in the translation of languages and bring people together. To give a programmed interpretation at such a large-scale, Facebook uses Machine Translation (MT) and, more specifically, Neural Machine Translation (NMT).

Machine Translation, a subfield of Computational Linguistics, utilizessoftware, tools, or programmes to translate written and verbal text from one language to another. MT mechanically replaces words from one language with words fromanother language, but that alone, often does not produce a decent translation. Moreover, not all the words in a single language have equal counterparts in another, numerous words have more than one translation., and two given dialects may have various structures.

Additionally, two given languages may have a completely different structure from one another. Tackling this issue with corpus, statistical and neural procedures is a quickly developing field that prompts better results and takes care of contrasts in linguistic typology, interpretation of idioms, and the segregation of anomalies.

Neural Machine Translation (NMT) is the leading approach to Machine Translation that produces improved translation by employing Artificial Neural Networks to calculate the possibility of word sequence (Sutskever, Vinyals, & Le, 2014). Facebook/ Metaverse is determined to enhance its feature of translation, for which purpose it has started using NMT. This study aims to analyze Facebook translation and find out errors and anomalies so that these translation tools can be further improved. To check the performance, Urdu to English translation will be tested. Urdu to English Facebook translation of 15 Facebook posts will be analyzed to check for any Neural Machine Translation (NMT) errors or anomalies.

Facebook/ Metaverse is widely used by people in Pakistan, and the majority of the population communicates using the Urdu language. This study will contribute to improving the translated text from Urdu to English, thus removing a language barrier in the way of effective communication and accurate news deliverance. Moreover, it will supplement the growth of the Urdu Language. Most importantly, it will aid the translation process and translation agencies.

Research Objectives

The ongoing study intends to realize these objectives:

i. To examine the competence of Neural Machine Translation (NMT).
ii. To highlight the errors produced by Facebook's Neural Machine Translation (NMT).

Research Questions

Keeping in view the research problem, this study raises the following questions:

i. What are the key shortcomings in Urdu to English translations of Facebook/ Metaverse?
ii. How do translation errors distort the semantic and syntactic structure in Urdu to English translation?

Literature Review

This section deals with studying and synthesizing recent relevant work done in this domain.

Furthermore, these studies established a premise for the present research.

In 2007, a study resented a method for machine translation of English to Urdu based on a transfer approach, case phrases, and verbs prepositions using the concepts of Paninnian grammar. Their Machine Translation design comprised of a syntactic-semantic approach. They were able to efficiently translate complex structures such as subordinate clauses with their design (Ata, Jawaid, & Kamran, 2007).

In 2013, Asad Abdul Malik and Asad Habib presented a comparative analysis of Example-Based Machine Translation, Rule-Based Machine Translation, and Statistical Machine Translation with their weaknesses and strengths. In 2013, Vaishali Gupta, Nisheeth Joshi and It Mathur presented a paper based on the Evaluation of English to Urdu Machine Translation based on two approaches, namely Automatic Evaluation (through tools) and Human Evaluation (through human cognition). The primary focus of the study remained on Human Evaluation which was used with diverse constraints to examine the quality of translated sentences. In this continuation, Gupta used MT systems like Google, Babylon, and Ijunoon, for automatic evaluation and metrics like BLEU, GTM, METEOR and

ATEC. He experimented on1000 corpus sentences. Moreover, he checked the correlation between human and automatic evaluation. All the metrics were used for English to Urdu machine-translated output. The results showed that the METEOR metric is greatly associated with humanly generated sentences (Gupta, Joshi, & Mathur, 2013).

In 2014, a seminal study discussed the challenges and problems encountered in English to Urdu Machine Translation. They classified these problems into four categories, given as follows:

a) Word Translation Problem.
b) Phrase Translation Problem.
c) Syntactic Translation Problem.
d) Semantic Translation Problem.

Furthermore, they proposed a parallel corpus as a key solution for these issues. Furthermore,they suggested alignment of the parallel corpus (Asnain, Wani, Arif, & Nazir, 2014).

In 2016, Sharmin Muzaffar, Pitambar Behera and Girish Nath Jha explored, classified, and resolved various types of divergence patterns in English-Urdu MT. One thousand corpora of the ILCI English sentences were taken for the experiment. Translated Urdu output through web-based MT platforms namely, Bing and Google Translate were analysed according to Dorr's (1994, 1995) theoretical framework (Muzaffar, Behera, & Jha, 2016).

In 2018, another study investigated translation with only large monolingual corpora available in both languages. The proposed neural and phrase-based models are simple and have fewer hyperparameters. In low resource languages such as English-Urdu, their method showed even better results (Lample, Ott, Conneau, Denoyer, & Ranzato, 2018).

In 2019, Usman Mahy ud Din highlighted the issues faced in out-of-vocabulary words (OOV) Machine Translation which inculcates named entities, technical terminologies, jargons, register, and foreign words. He presented the transliteration technique based on Expectation-Maximization (EM) which is un-supervised and language independent. He tested this technique on 3 models of SMT (Statistical Machine Translation), which included phrase-based, hierarchical phrase-based and factor-based models and two models of neural machine translation which included LSTM and transformer model. This research experimented with raw corpus and tokenized corpus and the results showed improvements of +3.5 points in baseline BLEU score (Ud Din, 2019). In 2019, a model was proposed for English to Urdu and Hindi Machine Translation systems using Translation Rules and Artificial Neural networks. The translation results derived from the system were evaluated using machine translation evaluationmethods. It has been observed that the system works efficiently on the trained linguistic (translation) rules and bilingual dictionaries. Therefore, an improvement in grammatical rules and the volume of the bilingual lexicon will result in an exact machine translation structure. The case marking is one of the significant features of the semantic accurateness of the translated manuscript. Case markers play their imperative roles in the transformation of meaning in Hindi and Urdu languages. The results showed that animprovement in case markings would produce better translations (Khan, & Usman, 2019).

In 2020, a study used English-Urdu machine translation to study four key areas including biomedical, theology, technical and General using Statistical as well as NMT. They checked several methods to decrease data dearth which includes, the use of programmed translation and transliterations. They performed several empirical studies to enhance the performance of each system and to explore the effect of data sources on the

classifications. Moreover, a comparison of the data sources and the effect of language model size on statistical machine translation performance was established (Abdul-Rauf, Abida, Zahra, Parvez, & Bashir, 2020).

In 2020, an approach related to removing gender and tense inconsistencies was presented in online English-Urdu translation systems. They used a Rule-Based Machine Translation Approach that supported Parts of Speech. Their model acquired 79% accuracy (Naeem, Salomi, & Khan, 2020). Transformers had been elongated towards trees in the process of neural machine translation. A few masks delimit nodes of trees. When the translation process is performed from English to German language or to other source languages, its accuracy enhances over seq to seq baseline, therefore, its results refine. (McDonald, & Chiang, 2021). Statistical and neural machine translation methods were adopted to translate from one language to the other. The deep learning method has been employed for translating the Urdu language into the English language and neural machine translation produces better results than a word to the word translation system. Two corpora in English and Urdu languages were built having 542810 and 540924 tokens respectively. Then the planned system was trained based on the 70:30 rule. Afterwards, it was juxtaposed with Google Translator to measure its efficacy, hence, it resulted in a 45.83 BLEU score (Andrabi, & Wahid, 2022).

The above-mentioned studies have contributed enormously to the field of Machine Translation by conducting experiments on websites like Google and Microsoft, but none of them focused on a prominent socializing platform like Facebook. Moreover, NMT is a novel approach to MT which needs to be tested for its accuracy. Thus, the current study is an expansion of the works in a different way.

Research Methodology

An exploratory study was directed to identify the problems existing in the Neural Machine Translator used by Facebook. Since this research is exploratory, the existing problems that are at a preliminary stage will be identified and analysed. This research is flexible and investigative and does not involve testing a hypothesis. It will provide a basis and aidfor future studies in this domain. The data was collected from the 15 selected Facebook posts by BBC Urdu/English and Voice of America Urdu/English, and then textually analysed to explore the semantic and syntactic errors made by the translator.

Research Design

This research focuses on the deficiencies in Urdu to English and English to Urdu translations in both English and Urdu versions of Facebook. Authentic news sources like BBC Urdu/English and Voice of America Urdu/English were selected for this research, and their posts were examined for grammatical, semantical, and syntactical errors in their translations. The mixed-methods approach was employed because it yields better results by combining both quantitative and qualitative data, methodologies, and paradigms in one research model. The key objective of this study is to investigate Facebook's Neural Machine Translation so that Facebook can enhance its tools according to the correct pattern of Urdu-English language and authentic lexicons,for example, The Oxford Urdu English Dictionary and FARHANG E AASFIA have been used as validation parameters.

Theoretical Framework

The selected phrases are analyzed based on the theoretical framework provided by Antonio Berman's "Twelve Deforming Tendencies". These tendencies are rationalization, clarification, expansion,

ennoblement and popularization, qualitative impoverishment, quantitative impoverishment, the destruction of rhythms, the destruction of underlying networks of signification, the destruction of linguistic patterning, the destruction of vernacular networks or their exoticization, the destruction of expressions or idioms and the effacement of the form.

Data Collection

Data was collected from selected 15 Facebook posts by 4 accounts of Urdu authors and playwrights namely Bano Qudsia, Khalil Ur Rehman Qamar, Umera Ahmed and Nemrah Ahmed. These posts are from 18th June 2016 to 10th November 2020. The data iscollected from Neural Translation of Facebook English and Urdu Captions. The generated results were checked for errors.

Results and Discussion

Following are the selected Urdu Facebook/ Metaverse posts which have been checked for their translation accuracy. The Oxford Urdu English Dictionary and FARHANG E AASFIA have been used as validation parameters. Each post has been analyzed in this section in detail through the lens of Antoine Berman's Twelve Deforming Tendencies.

Table 1: Urdu to English Translation

انسان کے اندر دو بھیڑیے ہوتے ہیں، ایک اچھائی کا، دوسرا برائی کا۔ غالب وہی رہتا ہے، جسے ہم کھلاتے پلاتے ہیں۔	Every human has two wolves, one for good and the other for evil. Ghalib is the one whom we feed and feed.
بتاؤ قیوم محبت کہاں ملتی ہے۔ کن کو ملتی ہے۔ میں اسے کیا بتاتا	Tell me where do you get Qayyum love, Who gets it, What should I tell him.
ہیی میرے پاس پچھتر پیسے ہیں۔ لیکن مجھے کوک پینا ہے۔ ہے کوئی اللہ تعالیٰ کا بندہ؟ اللہ تعالیٰ کا بندہ آفتاب ہمیشہ اُس کی ساتھ والی سیٹ پر ہوتا۔	Man, I have a pedestrian money but I have to drink coke. Is there any servant of ALLAH? The Servant of ALLAH, the sun is always on his seat.
نماز کی قضا ہے، لیکن خدمت کی قضا نہیں۔	There is a qaza of Prayer, but there is no qaza of service.
لفظوں کے دانت نہیں ہوتے مگر یہ کاٹ لیتے ہیں۔	Words don't have teeth but they cut it.
زندگی کا مقصد ذمہ داری ہے۔ اور سب سے بڑی ذمہ داری اللہ تعالیٰ کے ہاتھ میں ہاتھ دینا ہے۔	The purpose of life is responsibility, and the greatest responsibility is to give hands to ALLAH.
آپ کا لنگوٹیا یار جب آپ کی عزت کرنے لگے تو سمجھیں وہ آپ کی بے عزتی کر رہا ہے۔	When your langotia friend starts respecting you, then understand that he is insulting you.
ہر لڑکی جو پیار کرتی ہے بڑی کافر ہوتی ہے۔ بُت بنا کے رکھ لیتی ہے دل میں، خُدا سے چُھپ کر پوجتی ہے اُس کو اس لئے شاید بعد میں مجازی خُدا کہتے ہیں اُس کو۔	Every girl who loves is a very disbeliever, she makes an idol in her heart, and worship GOD secretly, that is why she may be called a virtual GOD Later.

ایسی باتیں کہاں سے سیکھی ہیں آپ نے ماما جان؟ کیا آپ نے چلے کاٹے ہیں؟ نہیں چلے نہیں کاٹے۔ میں نے غم بہت سہے ہیں۔ "غم کو صبر کے ساتھ سہنا چلا کاٹنے سے کم تو نہیں ہوتا لا حاصل"	From where did you learn such things, Uncle? Have you cut the walk? “Didn’t go, didn’t cut I’ve suffered a lot of sorrows. It’s nothing less than cutting the grief with patience Limitless.
دور دور تک پھیلے یہ وہ پہاڑ تھے جن کی پیشانیاں آسمان جھک کر چوم رہا تھا۔ وہ واقعی عظیم پہاڑ تھے اور ان کے درمیان میں قراقرم کا تاج محل کھڑا تھا جس کی سفید مرمریں دیواروں پر محبت کی ایک خاموش داستان لکھی تھی۔ وہ بلاشبہ آگرہ کے تاج محل سے زیادہ سفید اور حسین تھا۔	These are the mountains that spread far away, whose foreheads were kissing by the sky. They were really great mountains and the Taj Mahal of Karakoram was standing in the middle of them whose white marbles wrote a silent story of love on the walls. That of course the Taj Mahal of Agra was more white and beautiful that it was.
"میں بہارے گل ہوں۔ اناطولیہ کی بہارے گل" "تمہارا مطلب ہے گل بہار" نہیں! میں بہارے گل ہوں۔ یہ ایرانی نام ہے اور اس کا مطلب ہوتا ہے گلاب کے پھول پہ آئی بہار۔ پتا ہے میرا نام یہ کیوں ہے؟ "کیوں؟" کیونکہ میری آنم (ماں) کا نام آئے گل تھا۔ یعنی "چاند کا پھول"، میری نانی کا نام غنچے گل تھا اور میری بہن کا نام ہے عائشے گل۔ یعنی "وہ گلاب جو ہمیشہ زندہ رہے۔"	“I am bold girl. The Behare garden of Anatolia.” “You mean Gul Bahar” “No I am a bold girl. This is Iranian name and it means spring has come on rose flower. Do you know why name is this.” “Why” Because my mother’s name was Aaye Gul. That is the flower of the Moon, my grandmother’s name was Ghanche Gul and my Sister’s name is Ayesha Gul. Means the rose that will live Forever”
بہت دلچسپ۔ ترکی کے سارے پھول تو تمہارے خاندان میں ہیں۔ تمہارے بابا کا نام کیا ہوگا پھر؟ شاید گوبھی کا پھول۔	“Very interesting. All the flowers of Turkey are in your Family. What will be your Father’s name then, may be the cabbage flower.”
اُسامہ: "آپ کو وہ ویڈیو ان کے خلاف استعمال کرنی چاہئے تھی" سعدی: "یہ میرا طریقہ ہے اسے استعمال کرنے کا ہاشم کے خلاف۔ یقین کرو سیم، ہم اس کو ویسے استعمال نہیں کر سکتے تھے۔ ہر گیند کھیلنے والی نہیں ہوتی۔ کسی کسی گیند کو روکنا بھی ہوتا ہے۔ اُسامہ: "انسان کو کوئی چیز نہیں ہرا سکتی، جب تک کہ وہ خود ہار نہ مان لے۔ سعدی نے مشکوک نظروں سے اسے دیکھا۔ سعدی: "یہ کس کا ڈائیلاگ ہے؟" اُسامہ: "عمران خان کا ہے بھائی"	Osama: “You should have used that Video against them? Saadi: “This is my way of using in against Hashim. Believe me Sam” We couldn’t use it anyways. Not every ball is a player. Some balls has to be stopped. Osama: “Nothing can defeat a man until he himself accepts defeat. Saadi was him with suspicious eyes. Saad: “Whose dialogue is this” Osama: “This is Imran Khan’s Brother.”

اُس نے ہاتھ اٹھا کے ہولے سے نازک کانچ کی لڑی کو چھوا۔ وہ اسٹک سے ٹکرائی، اور لکڑی اور کانچ کی کوئی عجب سے دھن بج اُٹھی۔ موسیقی کی کسی بھی قسم سے مختلف، وہ کوئی انوکھی سے آواز تھی۔ اس کے لمس سے لڑیاں جو گول گول دائرے میں گھومنے لگی تھیں، اب آہستہ آہستہ ٹھہرنے کے قریب آرہی تھیں، اور تبھی اس نے دیکھا۔ اوپر کی سلور پلیٹ پہ انگریزی میں کھدا تھا۔	He raised her hand and touched the delicate glass fight. She hit the stick, and there was a strange tune of wood and glass. Different from any kind of music, she was a unique voice. From her touch, the Girls that were roaming around in the round circle, were now slowly approaching to stay, and then he saw. The silver plate on the top was dug in English
حیا! "کیا تم فارغ ہو؟" وہ بہت دوستانہ انداز میں پوچھ رہا تھا "ہاں! کیوں" اس نے دروازہ ذرا زیادہ کھول دیا تاکہ وہ بستر پر پھیلی اس کی کتابیں دیکھ کر جان لے کہ وہ ہرگز بھی فارغ نہیں ہے۔ اوکے! تم فارغ ہی ہو ٹھیک۔ "اُس نے سمجھ کر سر ہلایا"	"Shame on you! Are you free?" He was asking in a very friendly manner. "Yes! Why?" He opened the door a bit more so that he spreads on the bed, seeing his books and knowing that he is never free. "Okay! You are free, right. He thought it and made me cry.
"اور نہیں اس دنیا کی زندگی مگر سامان دھوکے کا۔"	"The life of this world is nothing but comfort of illusion."

The selected captions have deformed translation due to the shortcomings in different elements. The translation is getting damaged because of semantic, syntactic and transliteration errors. In the first example, Facebook has translated the Urdu word غالب into Roman English as "Ghalib" when the English translation for this would be "predominant". Moreover, the semantic meaning is deconstructed due to transliteration and the sentence has lost its meaning and sense completely. The translation of غالب as "Ghalib" syntactically deforms the sentence because the Urdu adjective is transliterated and then the first letter is capitalized making it appear like a proper noun. Therefore, this translation is entirely wrong as all the sense and meaning are lost in translation.

These errors reflect the deforming tendency of "rationalization" and "deconstruction of linguistic patterns" because the word order is destroyed due to the inaccuracy in translation.

The second Facebook NMT translation is also incorrect because the translation fails to preserve the correct sense and structure of the source language. Here the proper noun قیوم has been transliterated and damaged syntactically. The noun is changed into the adjective "Qayyum love"

when the correct translation for بتاؤ قیوم محبت کہاں ملتی ہے would have been, "Tell me Qayyum

where do you get love?" but instead the software-generated "Tell me where do you get Qayyumlove". The rest of the translation is carried out correctly, but the errors in this sentence negativelyaffects the meaning of the whole quotation. One possible reason for this error could the interpretation of قیوم محبت as an adjective but still, the word has not been translated into the target language. This results in the complete distortion of meaning and the sense. This aligns with the deforming tendency of "rationalization", "clarification" and "deconstruction of linguistic patterns".

In the third example, the correct translation is lost entirely, and the semantic meaning isdeconstructed. The Urdu number

پیدل is incorrectly translated as “pedestrian” and the expression becomes “pedestrian money” fails to convey the correct money. Moreover, the name آفتاب has been translated into its literal meaning which is “the sun”. Furthermore, سیٹ ساتھ والی is translated as “on this seat” which is wrong and delivers the opposite meaning. Therefore, the translation fails to reflect the correct meaning of the original text as the semantic and syntactic structure of the sentence is destroyed.

In the fourth and fifth posts, the generated translations do not make any sense and are therefore semantically wrong. In the fourth example, the Urdu word فضا has been transliterated because there is no accurate counterpart for it in the English language. The fifth example has an incorrect translation because یہ کاٹ لیتے ہیں۔ has been translated to “they cut it” so an extra object has been added which shows the deforming tendency of “expansion” because the sentence is unnecessarily prolonged. There is also an issue with the “clarification” because the correct meaning is not conveyed, and the original sense is lost.

The sixth translation is accurate at the beginning but the problem arises when the software deforms the syntactical structure by translating اللہ کے ہاتھ میں ہاتھ دینا to “give hands to Allah”which completely corrupts the original meaning of the text. This translation is evidence of the deforming tendency 'destruction of expressions and idioms' and 'qualitative impoverishment'. In the 7th instance, the translation is carried out correctly but the term لنگوٹیا has been transliterated when a corresponding translation for it already exists. It could have been translated as “childhoodpal” or “bosom friend” but the transliteration deforms the semantic sense and the intended meaning of the sentence is not efficiently conveyed. This shows the deforming tendency “destruction of linguistic patterns”.

The eighth post is an example of deforming tendency of "expansion" and "destruction of linguistic patterns". Here the translation lacks accuracy because of the damage done to the semantic and syntactic structure of the sentence which changes the meaning completely and the sentence fails to make any sense. بڑی کافر ہوتی ہے is translated as “is a very disbeliever” which is against the grammatical rules of English and the article “a” is an unnecessary addition to the sentence. Moreover, the pronoun "she" is added unnecessarily and the Urdu term خدا مجازی is translated inaccurately as "virtual God". These deformities contribute to deficiency in meaning and incorrect expressions.

The next selected posts are severely damaged in their semantically and syntactically. Here, "qualitative impoverishment" and "destruction of vernacular networks or their exoticization" are reflected as a deformity. ماما جان has been translated to "uncle" but it means "mama" or "mother" in English. ماما جان could also mean “maternal uncle” but the translation fails to convey that by simply translating it as “uncle”. Facebook NMT software fails to understand this word and mistranslates it. Furthermore, نہیں چلے نہیں کاٹے is interpreted incorrectly as "Didn't go Didn't cut" which renders the whole sentence meaningless. This is an example of the "destruction of vernacular networks or their exoticization" because taking out the vernacular of the prose has resulted in the eradication of cultural elements.

The tenth and eleventh example aligns with the deforming tendencies of “rationalization”, “clarification” and “destruction of linguistic patterns”. Here in these two examples, transliteration is done, and the quality of the vocabulary is damaged. The NMT were foreheads “whose is جن کی پیشانیاں for translation generated آسمان جھک کر چوم رہا تھا "kissing by the sky" instead of "whose

foreheads were being kissed by the sky". This results in an opposite meaning and the correct message of the sentence is destroyed. Moreover, the second error occurs in the translation of سے زیادہ سفید اور حسین تھا وہ بالشبہ آگرہ کے تاج محل which is translated as "That, of course, The Taj Mahal of Agra was whiter and beautiful than it was" which again damages the actual meaning and gives off the opposite one. The eleventh example has distorted punctuation along with various semantic and syntactic errors. The names have either been literally translated or transliterated. بہارے has first been transliterated as "bahare" and then incorrectly translated to "bold".

In the twelfth example, the sense of the sentence is lost in translation because ہر گیند کھیلنی والی نہیں ہوتی۔ is translated incorrectly as "not every ball is a player" which changes the original meaning and also does not make any semantic sense. Instead of using the past tense "played" the software has used the noun "player". In the thirteenth example, various deformities are evident which are distorting the sense of the translation. لڑیhere means "strand" but instead it has been interpreted as the verb لڑی and translated as "fight". Later, the plural of لڑی which is لڑیاں wrongly identified as لڑکیاں and then translated as "girls". ٹھہرنے کے قریب is literally translated as "approaching to stay" and کھدا is mistranslated as "dug" instead of "carved" because the context is ignored. This aligns with the deforming tendency of "clarification" and "destruction of linguistic patterns".

In the second last example, the mistranslated text renders the wrong meaning to the sentences. The proper noun حیاis translated as "shame on you" when it means "modesty". Moreover اس نے سمجھ کر سرہالیا is translated by the software as "he thought it and made me cry", not only is it an incorrect translation but also lacks any sense and meaning in the target language. This is an example of "destruction of linguistic patterns" and "clarification". In the last example, سامان is translated as "comfort" instead of the literal translation "luggage" which results in the correct translation of the sentence. One possible reason for this could be that this is a Quranic verse, and the software has proper translations for the Holy Quran in its directory. Therefore, the translation is semantically and syntactically completely accurate.

Conclusion

The abovementioned findings as a result of the analysis of the collected data provide the answer to both of our research questions as the errors have been individually pointed out and explained. The shortcomings in the translation of Urdu text to English text have been mentioned and their semantic and syntactic harm to the generated result has been explained. Furthermore, the deforming tendencies given by Berman have been identified in the data. The errors have been categorized and their frequency of occurrence has been observed. The most frequently occurring error was a semantic loss as the software failed to convey the intended meaning and the generated results made no sense. The second most occurring error was distorted syntactic pattern because the structure of the sentences was destroyed by translation and the word order and grammar were changed. Thirdly, transliteration was observed in various examples which deformed the sentences and prevented them from conveying the intended meaning. The results showed deforming tendencies of "rationalization", "clarification", "destruction of linguistic patterns", "destruction of idioms and expressions", "destruction of vernacular networks or their exoticization" and even "expansion" in some cases.

Facebook/ Metaverse is a widely used social networking site in Pakistan and the majority of its users communicate using Urdu language. These findings will help

Facebook improve its software which as a result will remove a language barrier in the way of effective communication. Moreover, it will aid the growth of Urdu as a language. Most importantly, it will help the translation process and translation agencies.

References

Abdul-Rauf, S., Abida, S., Zahra, S., Parvez, D., & Bashir, J. (2020, May). On the exploration of English to Urdu machine translation. In *Proceedings of the 1st Joint Workshop on Spoken Language Technologies for Under-resourced languages (SLTU) and Collaboration and Computing for Under-Resourced Languages (CCURL)* (pp. 285-293)

Andrabi, S. A. B., & Wahid, A. (2022). Machine Translation System Using Deep Learning for English to Urdu. *Computational Intelligence and Neuroscience*, *2022*.

Antoine, B. (2000). Translation and the Trials of the Foreign.

Asnain, S., Wani, N. A., Arif, T., & Nazir, M. (2014). An Analysis of Challenges in English to Urdu Machine Translation. *National Conference on Recent Innovations & Advancements in Information Technology*.

Ata, N., Jawaid, B., & Kamran, A. (2007). Rule based English to Urdu machine translation. In Proceedings of Conference on Language and Technology (CLT'07). Automatic-machine-translation-to-more-languages/

Bano Qudsia. (n.d.). Facebook - Log In or Sign Up. https://www.facebook.com/BanoQudsia

Dehlvi, M. S. A. (2002). Farhang-e-Asfia (ê ī h), reprinted by Sang-e.

Facebook Engineering. (2020, March 24). Expanding automatic machine translation to more languages. Retrieved from https://engineering.fb.com/ml-applications/expanding-

Facebook users worldwide 2020. (n.d.). Retrieved from https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-

Gupta, V., Joshi, N., & Mathur, I. (2013). Subjective and objective evaluation of English to Urdu Machine translation. In *2013 International Conference on Advances in Computing, Communications and Informatics (ICACCI)* (pp. 1520-1525). IEEE.

Qamar, K. U. R. (n.d.). Facebook - Log In or Sign Up. https://www.facebook.com/KrqOfficial1

Khan, S. N., & Usman, I. (2019). A model for English to Urdu and Hindi machine translation system using translation rules and artificial neural network. *Int. Arab J. Inf. Technol.*, *16*(1), 125-131.

Lample, G., Ott, M., Conneau, A., Denoyer, L., & Ranzato, M. A. (2018). Phrase-based & neural unsupervised machine translation. *arXiv preprint arXiv:1804.07755*.

McDonald, C., & Chiang, D. (2021, June). Syntax-Based Attention Masking for Neural Machine Translation. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Student Research Workshop* (pp. 47-52).

Muzaffar, S., Behera, P., & Jha, G. N. (2016). Classification and resolution of linguistic divergences in English-Urdu machine translation. *WILDRE: LREC*.

Nemrah, A. Official. (n.d.). Facebook - Log In or Sign Up. https://www.facebook.com/nemrah.ahmed.official proceedings of the 1st Joint SLTU and CCURL Workshop (SLTU-CCURL 2020) Language Resources and Evaluation Conference.

Sutskever, I., Vinyals, O., & Le, Q. V. (2014). Sequence to sequence learning with neural networks. *Advances in neural information processing systems*, *27*.

Thomas Fritz Albat. (2012). Systems and Methods for Automatically Estimating a Translation Time. Patent Application

Publication, 5.

Ud Din, U. M. (2019). Urdu-English Machine Transliteration using Neural Networks. https://doi.org/10.48550/arXiv.2001.05296

USA Today. (2016, September 29). Does Facebook speak your language? Retrieved from https://www.usatoday.com/story/tech/news/2016/09/29/facebook-translation-new- languages-corsican-fulah,malta/91268284/#:~:text=In%20all%2C%20Facebook%

Venuti, L. (2012). The translation studies reader. London: Routledge