World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:18, No:11, 2024

ALEF: An Enhanced Approach to Arabic-English Bilingual Translation

Authors : Abdul Muqsit Abbasi, Ibrahim Chhipa, Asad Anwer, Saad Farooq, Hassan Berry, Sonu Kumar, Sundar Ali, Muhammad Owais Mahmood, Areeb Ur Rehman, Bahram Baloch

Abstract : Accurate translation between structurally diverse languages, such as Arabic and English, presents a critical challenge in natural language processing due to significant linguistic and cultural differences. This paper investigates the effectiveness of Facebook's mBART model, fine-tuned specifically for sequence-tosequence (seq2seq) translation tasks between Arabic and English, and enhanced through advanced refinement techniques. Our approach leverages the Alef Dataset, a meticulously curated parallel corpus spanning various domains to capture the linguistic richness, nuances, and contextual accuracy essential for high-quality translation. We further refine the model's output using advanced language models such as GPT-3.5 and GPT-4, which improve fluency, coherence, and correct grammatical errors in translated texts. The fine-tuned model demonstrates substantial improvements, achieving a BLEU score of 38.97, METEOR score of 58.11, and TER score of 56.33, surpassing widely used systems such as Google Translate. These results underscore the potential of mBART, combined with refinement strategies, to bridge the translation gap between Arabic and English, providing a reliable, context-aware machine translation solution that is robust across diverse linguistic contexts.

Keywords : natural language processing, machine translation, fine-tuning, Arabic-English translation, transformer models, seq2seq translation, translation evaluation metrics, cross-linguistic communication
Conference Title : ICMLC 2024 : International Conference on Machine Learning and Cybernetics
Conference Location : London, United Kingdom
Conference Dates : November 25-26, 2024