Enhancing Word Meaning Retrieval Using FastText and Natural Language Processing Techniques

Authors : Sankalp Devanand, Prateek Agasimani, Shamith V. S., Rohith Neeraje

Abstract : Machine translation has witnessed significant advancements in recent years, but the translation of languages with distinct linguistic characteristics, such as English and Sanskrit, remains a challenging task. This research presents the development of a dedicated English-to-Sanskrit machine translation model, aiming to bridge the linguistic and cultural gap between these two languages. Using a variety of natural language processing (NLP) approaches, including FastText embeddings, this research proposes a thorough method to improve word meaning retrieval. Data preparation, part-of-speech tagging, dictionary searches, and transliteration are all included in the methodology. The study also addresses the implementation of an interpreter pattern and uses a word similarity task to assess the quality of word embeddings. The experimental outcomes show how the suggested approach may be used to enhance word meaning retrieval tasks with greater efficacy, accuracy, and adaptability. Evaluation of the model's performance is conducted through rigorous testing, comparing its output against existing machine translation systems. The assessment includes quantitative metrics such as BLEU scores, METEOR scores, Jaccard Similarity, etc.