

Abstract

PraTranV3 is a machine translation project focused on translating Ardha Magadhi Prakrit, an ancient Middle Indo-Aryan language, into English. The project addresses the challenges posed by the scarcity of digitized resources, OCR inaccuracies, and the linguistic complexities of ancient languages. Building upon previous work in Maharashtri Prakrit translation, this system leverages state-of-the-art neural machine translation (NMT) models, such as LLAMA, Claude, NLLB, and M2M100, fine-tuned with domain-specific data from ancient manuscripts. The project aims to preserve cultural heritage by providing a web-based translation tool accessible to researchers, historians, and the public. The outcomes include the development of a functional translation system, contributions to low-resource language processing, and a scalable framework for other ancient languages.

Keywords: Ardha Magadhi Prakrit, Machine Translation, Neural Machine Translation (NMT), Low-Resource Languages, OCR, Data Preprocessing, Fine-Tuning, LLAMA, Claude, NLLB, M2M100, Cultural Preservation, Digital Heritage, Web Application.