Abstract

In this paper, it is described how an Assamese sentence is translated to English using statistical machine translation. Statistical Machine Translation is the paradigm where translations from source to target language are based on statistical models. Moses is used as a platform for Statistical Machine Translation. GIZA++ is also used for word-alignment and IRSTLM for language model training. A Transliteration model is also integrated into the system to deal with out of vocabulary (OOV) words.

References

- Marian Olteanu et al, “Phramer: Open Source Statistical Phrase-Based Translator”, In the Proceedings of the Workshop of Statistical Machine Translation, June
2006, pp. 146-149.
- Philipp Koehn et al., "Statistical Phrase-Based Translation"; In the Proceedings of HLT-NAALC, May-June 2003, pp. 48-54.
- "PSMT"; Available: http://psmt.sourceforge.net/

Index Terms

Computer Science Artificial Intelligence

Keywords

Assamese English Statistical Machine Translation Transliteration Corpus.