Abstract

Named Entity Recognition is a process through which a program extracts proper nouns in texts and associates them with a proper tag. NER has made significant progress in European languages, but in Indian languages due to the lack of effort as well as proper resources, it remains a challenging task. Recognizing ambiguities and assigning the correct tags to the names is the main goal of NER. Thus NER can be defined as identification of the proper nouns and the classification of these nouns into classes such as person, location, organization and miscellaneous including date, time and year. The main aim of this work is to develop a computational system that can perform NER in text in Assamese, which is a resource poor Indo-Aryan language. This article presents an overview of NER and its issues in the context of Assamese, and also the work done in Assamese using various approaches.

References

Science Dept, New York University, 1999.


Index Terms

Computer Science Information Sciences
Keywords

Named Entity Recognition, Assamese Corpora