Survey of Named Entity Recognition Techniques for Various Indian Regional Languages

Foundational Journal of Computer Applications
Foundation of Computer Science (FCS), NY, USA

Volume 164
Number 4

Year of Publication: 2017

Authors:

Shrutika Kale, Sharvari Govilkar

10.5120/ijca2017913621

Abstract

Named entity recognition is a process and study of identification of entities that are proper nouns and classifying them to their appropriate pre-defined class, also called as tag. Named entity recognition is also called as entity chunking, entity identification and entity extraction. It is a sub task of information extraction, where structured text is extracted from unstructured text. Popular applications of NER are machine translation, text mining, data classification, question answering system. This paper presents a survey of different NER techniques, approach, observations and features for Indian regional languages such as Hindi, Marathi, Bengali, Punjabi, Malayalam, Bengali, Kannada, Telugu, Tamil, Urdu, Oriya.

References

Survey of Named Entity Recognition Techniques for Various Indian Regional Languages

Model (HMM): An Experimental Result on Hindi, Urdu and Marathi Languages”, April 2013.

Index Terms

Computer Science Pattern Recognition

Keywords

Named entity recognition, Named Entity, Machine learning techniques, Statistical Model, NERC