Abstract

Part of Speech tagging has a vital role in different fields of natural language processing. It can be defined as the process of assigning a tag or a label to a word according to its morphological or syntactical properties. The objective of this paper is to develop a POS tagger based on hybrid approach which is combination of rule based approach and CRF based approach. In this, the tagset used 36 tags which is proposed by TDIL for Indian languages.

References

Assigning the Correct Word Class to Punjabi Unknown Words using CRF

(ICCACL), 978- 1-4799-3080-71 14/$31.00 ©20 14 IEEE.
4. Garg Navneet, Goyal Vishal, Suman Preet,” Rule Based Hindi Part of Speech Tagger”,
Proceedings of COLING 2012: Demonstration Papers, pages 163–174, COLING 2012,
Mumbai, December 2012.
5. Joshi Nisheeth, Darbari Hemant, Mathur Iti, “HMM based POS tagger for Hindi”, Jan
Indian Languages: A Survey”, International Journal of Computer Applications (0975 – 8887),
Volume 6– No.5, September 2010.
Of Punjabi Unknown Words”, IMS - International Conference on Information and Mathematical
Sciences 2013.
Language in India, www. languageinindia.com, Strength for Today and Bright Hope for
Tomorrow“, Volume 9: 10 October 2009, ISSN 1930-2940.
Accuracy of Punjabi POS Tagger”, 978-1-4244-8728-8/11/$26.00 ©2011IEEE.
11. Manchanda Blossom, Ravishanker,” To find the POS tag of unknown words in Punjabi
language”, An International Journal of Engineering Sciences ISSN: 2229-6913 Issue July 2011,
Vol. 1.
12. Mohnot Kanak, Bansal Neha, Singh Shashi Pal, Kumar Ajai, “Hybrid approach for Part
of Speech Tagger for Hindi language”, International Journal of Computer Technology and
Electronics Engineering (IJCTEE), Volume 4, Issue 1, February 2014.
processing: An introduction", J Am Med Inform Assoc 2011; 18:544e551. DOI:
10.1136/amiajn-2011-000464, Published by group.bmj.com on October 5, 2011.
Random Fields”, Proceedings of the IJCNLP-08 Workshop on NLP for Less Privileged
Languages, pages 117–122, Hyderabad, India, January 2008. Asian Federation of Natural
Language Processing.
15. Singh Thoudam Doren, Ekbal Asif, Bandypadhyay Sivaji. “Manipuri POS Tagging using
CRF and SVM: A Language Independent Approach”, 6th International Conference on Natural
Language Processing, 2008.
16. Singha Kh Raju, Purkayastha Bipul Syam, Singha Kh Dhiren, “Part of Speech Tagging in
8887), Volume 51– No.14, August 2012.
17. V Krishnapriya, P Sreesha, T.R. Harithalakshmi, T.C. Archana, Vettath Jayasree N,
“Design of a POS Tagger using Conditional Random Fields for Malayalam”, First International
Conference on Computational Systems and Communications (ICCSC), Trivandrum,
978-1-4799-6013-2/14/$31.00 ©2014 IEEE.
18. [http://tdil.mit.gov.in/

Index Terms

Computer Science    Pattern Recognition
Keywords

Natural Language Processing, Part of Speech Tagging, Rule based approach, CRF, Hybrid.