

12th International Conference on Natural Language Processing (ICON-2015)

Trivandrum, India
11 – 14 December 2015

ISBN: 978-1-5108-3787-4

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571



Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2015) by the Association for Computational Linguistics
All rights reserved.

Printed by Curran Associates, Inc. (2017)

For permission requests, please contact the Association for Computational Linguistics
at the address below.

Association for Computational Linguistics
209 N. Eighth Street
Stroudsburg, Pennsylvania 18360

Phone: 1-570-476-8006
Fax: 1-570-476-0860

acl@aclweb.org

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2633
Email: curran@proceedings.com
Web: www.proceedings.com

Table of Contents

<i>Keynote Lecture 1: Scientific Paper Analysis</i>	
Yuji Matsumoto	
<i>Addressing Class Imbalance in Grammatical Error Detection with Evaluation Metric Optimization</i>	
Anoop Kunchukuttan and Pushpak Bhattacharyya	
<i>Words are not Equal: Graded Weighting Model for Building Composite Document Vectors</i>	
Pranjal Singh and Amitabha Mukerjee	
<i>Online Adspace Posts' Category Classification</i>	
Dhawal Joharapurkar, Vaishak Salin and Vishal Krishna	
<i>Noun Phrase Chunking for Marathi using Distant Supervision</i>	
Sachin Pawar, Nitin Ramrakhiani, Girish K. Palshikar, Pushpak Bhattacharyya and Swapnil Hingmire	
<i>Self-Organizing Maps for Classification of a Multi-Labeled Corpus</i>	
Lars Bungum and Björn Gambäck	
<i>Word Sense Disambiguation in Hindi Language Using Hyperspace Analogue to Language and Fuzzy C-Means Clustering</i>	
Devendra K. Tayal, Leena Ahuja and Shreya Chhabra	
<i>Using Word Embeddings for Bilingual Unsupervised WSD</i>	
Sudha Bhingardive, Dharendra Singh, Rudramurthy V and Pushpak Bhattacharyya	
<i>Compositionality in Bangla Compound Verbs and their Processing in the Mental Lexicon</i>	
Tirthankar Dasgupta, Manjira Sinha and Anupam Basu	
<i>IndoWordNet Dictionary: An Online Multilingual Dictionary using IndoWordNet</i>	
Hanumant Redkar, Sandhya Singh, Nilesh Joshi, Anupam Ghosh and Pushpak Bhattacharyya ..	
<i>Let Sense Bags Do Talking: Cross Lingual Word Semantic Similarity for English and Hindi</i>	
Apurva Nagvenkar, Jyoti Pawar and Pushpak Bhattacharyya	
<i>A temporal expression recognition system for medical documents by</i>	
Naman Gupta, Aditya Joshi and Pushpak Bhattacharyya	
<i>An unsupervised EM method to infer time variation in sense probabilities</i>	
Martin Emms and Arun Jayapal	
<i>Solving Data Sparsity by Morphology Injection in Factored SMT</i>	
Sreelekha S, Piyush Dungarwal, Pushpak Bhattacharyya and Malathi D	
<i>Authorship Attribution in Bengali Language</i>	
Shanta Phani, Shibamouli Lahiri and Arindam Biswas	

<i>TransChat: Cross-Lingual Instant Messaging for Indian Languages</i>	Diptesh Kanojia, Shehzaad Dhuliawala, Abhijit Mishra, Naman Gupta and Pushpak Bhattacharyya
<i>A Database of Infant Cry Sounds to Study the Likely Cause of Cry</i>	Shivam Sharma, Shubham Asthana and V. K. Mittal
<i>Perplexed Bayes Classifier</i>	Cohan Sujay Carlos.....
<i>An Empirical Study of Diversity of Word Alignment and its Symmetrization Techniques for System Combination</i>	Thoudam Doren Singh
<i>Domain Sentiment Matters: A Two Stage Sentiment Analyzer</i>	Raksha Sharma and Pushpak Bhattacharyya.....
<i>Extracting Information from Indian First Names</i>	Akshay Gulati.....
<i>punct-An Alternative Verb Semantic Ontology Representation</i>	Kavitha Rajan.....
<i>SMT Errors Requiring Grammatical Knowledge for Prevention</i>	Yukiko Sasaki Alam
<i>Isolated Word Recognition System for Malayalam using Machine Learning</i>	Maya Moneykumar, Elizabeth Sherly and Win Sam Varghese
<i>Judge a Book by its Cover: Conservative Focused Crawling under Resource Constraints</i>	Shehzaad Dhuliawala, Arjun Atreya V, Ravi Kumar Yadav and Pushpak Bhattacharyya.....
<i>Text Normalization and Unit Selection for a Memory Based Non Uniform Unit Selection TTS in Malayalam</i>	Gokul P., Neethu Thomas, Crisil Thomas and Dr. Deepa P. Gopinath
<i>Morphological Analyzer for Gujarati using Paradigm based approach with Knowledge based and Statistical Methods</i>	Jatayu Baxi, Pooja Patel and Brijesh Bhatt
<i>Resolution of Pronominal Anaphora for Telugu Dialogues</i>	Hemanth Reddy Jonnalagadda and Radhika Mamidi
<i>A Study on Divergence in Malayalam and Tamil Language in Machine Translation Perceptive</i>	Jisha P Jayan and Elizabeth Sherly
<i>Automatic conversion of Indian Language Morphological Processors into Grammatical Framework (GF)</i>	Harsha Vardhan Grandhi and Soma Paul

<i>Logistic Regression for Automatic Lexical Level Morphological Paradigm Selection for Konkani Nouns</i>	Shilpa Desai, Jyoti Pawar and Pushpak Bhattacharyya.....
<i>Ruchi: Rating Individual Food Items in Restaurant Reviews</i>	Burusothman Ahiladas, Paraneetharan Saravanaperumal, Sanjith Balachandran, Thamayanthy Sri- palan and Surangika Ranathunga
<i>Dependency Extraction for Knowledge-based Domain Classification</i>	Lokesh Kumar Sharma and Namita Mittal
<i>An Approach to Collective Entity Linking</i>	Ashish Kulkarni, Kanika Agarwal, pararth Shah, Sunny Raj Rathod and Ganesh Ramakrishnan
<i>Development of Speech corpora for different Speech Recognition tasks in Malayalam language</i>	Cini Kurian
<i>POS Tagging of Hindi-English Code Mixed Text from Social Media: Some Machine Learning Experi- ments</i>	Royal Sequiera, Monojit Choudhury and Kalika Bali.....
<i>Automated Analysis of Bangla Poetry for Classification and Poet Identification</i>	Geetanjali Rakshit, Anupam Ghosh, Pushpak Bhattacharyya and Gholamreza Haffari
<i>Sentence Boundary Detection for Social Media Text</i>	Dwijen Rudrapal, Anupam Jamatia, Kunal Chakma, Amitava Das and Björn Gambäck
<i>Mood Classification of Hindi Songs based on Lyrics</i>	Braja Gopal Patra, Dipankar Das and Sivaji Bandyopadhyay
<i>Using Skipgrams, Bigrams, and Part of Speech Features for Sentiment Classification of Twitter Mes- sages</i>	Badr Mohammed Badr and S. Sameen Fatima
<i>A Hybrid Approach for Bracketing Noun Sequence</i>	Arpita Batra and Soma Paul
<i>Simultaneous Feature Selection and Parameter Optimization Using Multi-objective Optimization for Sentiment Analysis</i>	Mohammed Arif Khan, Asif Ekbal and Eneldo Loza Mencía.....
<i>Detection of Multiword Expressions for Hindi Language using Word Embeddings and WordNet-based Features</i>	Dhirendra Singh, Sudha Bhingardive, Kevin Patel and Pushpak Bhattacharyya.....
<i>Augmenting Pivot based SMT with word segmentation</i>	Rohit More, Anoop Kunchukuttan, Pushpak Bhattacharyya and Raj Dabre
<i>Using Multilingual Topic Models for Improved Alignment in English-Hindi MT</i>	Diptesh Kanojia, Aditya Joshi, Pushpak Bhattacharyya and Mark James Carman.....

Triangulation of Reordering Tables: An Advancement Over Phrase Table Triangulation in Pivot-Based SMT

Deepak Patil, Harshad Chavan and Pushpak Bhattacharyya

Post-editing a chapter of a specialized textbook into 7 languages: importance of terminological proximity with English for productivity

Ritesh Shah, Christian Boitet, Pushpak Bhattacharyya, Mithun Padmakumar, Leonardo Zilio, Ruslan Kalitvianski, Mohammad Nasiruddin, Mutsuko Tomokiyo and Sandra Castellanos Páez

Generating Translation Corpora in Indic Languages: Cultivating Bilingual Texts for Cross Lingual Fertilization

Niladri Sekhar Dash, Arulmozi Selvraj and Mazhar Hussain

Translation Quality and Effort: Options versus Post-editing

Donald Sturgeon and John S. Y. Lee

Investigating the potential of post-ordering SMT output to improve translation quality

Pratik Mehta, Anoop Kunchukuttan and Pushpak Bhattacharyya

Applying Sanskrit Concepts for Reordering in MT

Akshar Bharati, , Prajna Jha, Soma Paul and Dipti M Sharma

Dialogue Act Recognition for Text-based Sinhala

Sudheera Palihakkara, Dammina Sahabandu, Ahsan Shamsudeen, Chamika Bandara and Surangika Ranathunga

A Semi Supervised Dialog Act Tagging for Telugu

Suman Dowlagar and Radhika Mamidi

Ranking Model with a Reduced Feature Set for an Automated Question Generation System

Manisha Satish Divate and Ambuja Salgaonkar

Natural Language Processing for Solving Simple Word Problems

Sowmya S Sundaram and Deepak Khemani

Analysis of Influence of L2 English Speakers' Fluency on Occurrence and Duration of Sentence-medial Pauses in English Readout Speech

Shambhu Nath Saha and Shyamal Kr. Das Mandal

Acoustic Correlates of Voicing and Gemination in Bangla

Aanusha Ghosh