Preface

Research in Natural Language Processing (NLP) has taken a noticeable leap in the recent years. Tremendous growth of information on the web and its easy access has stimulated large interest in the field. India with multiple languages and continuous growth of Indian language content on the web makes a fertile ground for NLP research. Moreover, industry is keenly interested in obtaining NLP technology for mass use. The internet search companies are increasingly aware of the large market for processing languages other than English. For example, search capability is needed for content in Indian and other languages. There is also a need for searching content in multiple languages, and making the retrieved documents available in the language of the user. As a result, a strong need is being felt for machine translation to handle this large instantaneous use. Information Extraction, Question Answering Systems and Sentiment Analysis are also showing up as other business opportunities.

These needs have resulted in two welcome trends. First, there is much wider student interest in getting into NLP at both postgraduate and undergraduate levels. Many students interested in computing technology are getting interested in natural language technology, and those interested in pursuing computing research are joining NLP research. Second, the research community in academic institutions and the government funding agencies in India have joined hands to launch consortia projects to develop NLP products. Each consortium project is a multi-institutional endeavour working with a common software framework, common language standards, and common technology engines for all the different languages covered in the consortium. As a result, it has already led to development of basic tools for multiple languages which are inter-operable for machine translation, cross lingual search, hand writing recognition and OCR.

In this backdrop of increased student interest, greater funding and most importantly, common standards and interoperable tools, there has been a spurt in research in NLP on Indian languages whose effects we have just begun to see. A great number of submissions reflecting good research is a heartening matter. There is an increasing realization to take advantage of features common to Indian languages in machine learning. It is a delight to see that such features are not just specific to Indian languages but to a large number of languages of the world, hitherto ignored. The insights so gained are furthering our linguistic understanding and will help in technology development for hopefully all languages of the world.

For machine learning and other purposes, linguistically annotated corpora using the common standards have become available for multiple Indian languages. They have been used for the development of basic technologies for several languages. Larger set of corpora are expected to be prepared in near future.

This conference proceedings contains papers selected for presentation in technical sessions of ICON-2017 and short communications selected for poster presentation. We are thankful to our excellent team of reviewers from all over the globe who deserve full credit for the hard work of reviewing the high quality submissions with rich technical content. From 141 submissions, 64 papers were selected, 32 for full presentation, 32 for poster presentation, representing a variety of new and interesting developments, covering a wide spectrum of NLP areas and core linguistics.

We are deeply grateful to Bjørn W. Schuller, University of Passau, Germany, NG Hwee Tou, National University of Singapore (NUS), Singapore and Vasudeva Varma, IIIT Hyderabad, India for giving the keynote lectures at ICON. We would also like to thank the members of the Advisory Committee and
Programme Committee for their support and co-operation in making ICON 2017 a success.

We thank Anil Kumar Singh, Chair, Student Paper Competition and Dipankar Das, Chair, NLP Tools Contest for taking the responsibilities of the events. We are thankful to Sudip Kumar Naskar and Dipankar Das for making the organization of the event at Jadavpur University a success.

We convey our thanks to P V S Ram Babu, G Srinivas Rao, B Mahender Kumar and A Lakshmi Narayana, International Institute of Information Technology (IIIT), Hyderabad for their dedicated efforts in successfully handling the ICON Secretariat. We also thank IIIT Hyderabad team of Vineet Chaitanya, Vasudeva Varma, Soma Paul, Radhika Mamidi, Manish Shrivastava, Suryakanth V Gangashetty and Anil Kumar Vuppala. We heartfully express our gratitude to Somnath Banerjee, Tapabrata Mondal, Sainik Mahata and other team members at Jadavpur University for their timely help with sincere dedication to make this conference a success.

We also thank all those who came forward to help us in this task.

Finally, we thank all the researchers who responded to our call for papers and all the participants of ICON-2017, without whose overwhelming response the conference would not have been a success.

December 2017
Varanasi

Sivaji Bandyopadhyay
Dipti Misra Sharma
Rajeev Sangal
Advisory Committee:

Aravind K Joshi, University of Pennsylvania, USA (Chair)

Conference General Chair:

Rajeev Sangal, IIT (BHU), Varanasi, India

Programme Committee:

Sivaji Bandyopadhyay, Jadavpur University,Kolkata, India (Chair)  
Dipti Misra Sharma, IIT Hyderabad, India (Co-Chair) 
Kalika Bali, Microsoft Research India, India  
Srinivas Bangalore, Interactions LLC, AT&T Research, USA  
Rajesh Bhatt, University of Massachusetts, USA  
Pushpak Bhattacharyya, IIT Patna, India  
Monojit Choudhury, Microsoft Research India, India  
Josef van Genabith, DFKI GmbH, Germany  
Harald Hamrarstrm, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands  
Mohammed Hasanuzzaman, Universite de Caen, Normandie, France  
Gurpreet Lehal, Punjabi University, Patiala, India  
Yuji Matsumoto, Nara Institute of Science and Technology, Japan  
Joakim Nivre, Uppsala University, Sweden  
Alexis Palmer, Heidelberg University, Germany  
Jyoti Pawar, DCST, Goa University, India  
Owen Rambow, University of Columbia, USA  
Paolo Rosso, Universitat Politcnica de Valencia, Spain  
Shikhar K. Sarma, Gauhati University, India  
Elizabeth Sherly, IIITM-K, Trivandrum, India  
Shobha Lalitha Devi, AU-KBC, Chennai, India  
Keh-Yih Su, Institute of Information Science, Academia Sinica, Taiwan  
Vasudeva Varma, IIIT Hyderabad, India

Tools Contest Chairs:

Dipankar Das, Jadavpur University, Kolkata, India (Chair)  
Amitava Das, IIIT-Sri City, India

Student Paper Competition Chair:

Anil Kumar Singh, IIT (BHU), India

Organizing Committee:

Sudip Naskar, Jadavpur University, Kolkata, India  
Dipankar Das, Jadavpur University, Kolkata, India
We gratefully acknowledge the excellent quality of refereeing we received from the reviewers. We thank them all for being precise and fair in their assessment and for reviewing the papers in time.

Abhijeet Mishra
Adithya Pratapa
Aditi Mukherjee
Aditya Joshi
Aishwarya N Reganti
Akhilesh Sudhakar
Alok Ranjan Pal
Amitava Das
Amol Bole
Anand Kumar
Anchal Rani
Anil Kumar Singh
Anil Thakur
Anil Kumar Vuppala
Animesh Mukherjee
Anshul Bawa
Anupam Jamatia
Anupam Mondal
Anusha Prakash
Anushiya Rachel G
Arjun Akula
Arun Baby
Ashika Naidu
Ashwini Vaidya
Asif Ekbal
Aswin Shanmugam S
Bajibabu Bollepalli
Balamurali A R
Bapi Raju
Basil Abraham
Bharat Ram Ambati
Bhuvana Narasimhan
Braja Gopal Patra
C V Jawahar
Chandra Sekhar Chellu
Christian Boitet
D S Karthik Pandia
Dan Zeman
Deepak P
Delia Irazú Hernández Farias
Devi G
Dipankar Das
Dipti Misra Sharma

Divya Sai Jitta
Dwijen Rudrapal
Elizabeth Sherly
Erik Cambria
Fei Xia
Girish Palshikar
Golda Brunet Rajan
Gurpreet Singh Lehal
Harald Hammarström
Hema Murthy
Himanshu Sharma
Irshad Bhat
Jilt Sebastian
Jose Moreno
Joy Mahapatra
Jyoti Pawar
Kalika Bali
Kamal Sarkar
Karunesh Arora
Keh-Yih Su
Kevin Patel
Kishorjit Nongmeikapam
Kunal Chakma
Litton J Kurisinkel
Mahidas Bhattacharya
Malhar Kulkarni
Manish Shrivastava
Manoj Chinnakotla
Maria Anzovino
Mayank Singh
Mohammed Hasanuzzaman
Monojit Choudhury
Nikhil Pattisapu
Nilooofar Safi
Niraj Kumar
Nirmal Surange
Owen Rambow
Paolo Rosso
Parameswari Krishnamurthy
Partha Pakray
Partha Talukdar
Pawan Goyal
Pranav Dhakras
# Table of Contents

**Keynote Lecture 1: NLP in Tomorrow’s Profiling - Words May Fail You**  
Björn W. Schuller ............................................................ 1

**Deriving Word Prosody from Orthography in Hindi**  
Somnath Roy ........................................................................ 2

**Three-phase training to address data sparsity in Neural Machine Translation**  
Ruchit Agrawal, Mihir Shekhar and Dipti Sharma ..................... 13

**Reference Scope Identification for Citances Using Convolutional Neural Networks**  
Saurav Jha, Aanchal Chaurasia, Akhilesh Sudhakar and Anil Kumar Singh .......... 23

**A vis-à-vis evaluation of MT paradigms for linguistically distant languages**  
Ruchit Agrawal, Jahfar Ali and Dipti Misra Sharma ..................... 33

**Textual Relations and Topic-Projection: Issues in Text Categorization**  
Lahari Chatterjee, Samir Karmakar and Abahan Datta ................. 43

**POS Tagging For Resource Poor Languages Through Feature Projection**  
Pruthwik Mishra, Vandan Mujadia and Dipti Misra Sharma ............. 50

**An Exploration of Word Embedding Initialization in Deep-Learning Tasks**  
Tom Kocmi and Ondrej Bojar ................................................. 56

**Curriculum Design for Code-switching: Experiments with Language Identification and Language Modeling with Deep Neural Networks**  
Monojit Choudhury, Kalika Bali, Sunayana Sitaram and Ashutosh Baheti .......... 65

**Quantitative Characterization of Code Switching Patterns in Complex Multi-Party Conversations: A Case Study on Hindi Movie Scripts**  
Adithya Pratapa and Monojit Choudhury ................................... 75

**Towards Normalising Konkani-English Code-Mixed Social Media Text**  
Akshata Phadte and Gaurish Thakkar ........................................ 85

**Towards developing a phonetically balanced code-mixed speech corpus for Hindi-English ASR**  
Ayushi Pandey, Brij Mohan Lal Srivastava and Suryakanth Gangashetty .......... 95

**Keynote Lecture 2: Grammatical Error Correction: Past, Present and Future**  
NG Hwee Tou .................................................................. 102

**Hybrid Approach for Marathi Named Entity Recognition**  
Nita Patil, Ajay Patil and B.V. Pawar ....................................... 103

**Sentiment Analysis: An Empirical Comparative Study of Various Machine Learning Approaches**  
Swapnil Jain, Shrikant Malviya, Rohit Mishra and Uma Shanker Tiwary .......... 112
<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Handling Multi-Sentence Queries in a Domain Independent Dialogue System</td>
<td>Prathyusha Jwalapuram and Radhika Mamidi</td>
<td>122</td>
</tr>
<tr>
<td>Document Level Novelty Detection: Textual Entailment Lends a Helping Hand</td>
<td>Tanik Saikh, Tirthankar Ghosal, Asif Ekbak and Pushpak Bhattacharyya</td>
<td>131</td>
</tr>
<tr>
<td>Is your Statement Purposeless? Predicting Computer Science Graduation Admission Acceptance based on Statement Of Purpose</td>
<td>Diptesh Kanojia, Nikhil Wani and Pushpak Bhattacharyya</td>
<td>141</td>
</tr>
<tr>
<td>Natural Language Programing with Automatic Code Generation towards Solving Addition-Subtraction Word Problems</td>
<td>Sourav Mandal and Sudip Kumar Naskar</td>
<td>146</td>
</tr>
<tr>
<td>Unsupervised Separation of Transliterable and Native Words for Malayalam</td>
<td>Deepak P</td>
<td>155</td>
</tr>
<tr>
<td>Known Strangers: Cross Linguistic Patterns in Multilingual Multidirectional Dictionaries</td>
<td>Rejitha K. S. and Rajesha N.</td>
<td>165</td>
</tr>
<tr>
<td>Tutorial for Deaf Teaching Punjabi Alphabet using Synthetic Animations</td>
<td>Lalit Goyal and Vishal Goyal</td>
<td>172</td>
</tr>
<tr>
<td>SemTagger: A Novel Approach for Semantic Similarity Based Hashtag Recommendation on Twitter</td>
<td>Kuntal Dey, Ritvik Shrivastava, Saroj Kaushik and L. Venkata Subramaniam</td>
<td>178</td>
</tr>
<tr>
<td>Reasoning with Sets to Solve Simple Word Problems Automatically</td>
<td>Sowmya S Sundaram and Deepak Khemani</td>
<td>188</td>
</tr>
<tr>
<td>Improving NER for Clinical Texts by Ensemble Approach using Segment Representations</td>
<td>Hamada Nayel and H L Shashirekha</td>
<td>197</td>
</tr>
<tr>
<td>Beyond Word2Vec: Embedding Words and Phrases in Same Vector Space</td>
<td>Vijay Prakash Dwivedi and Manish Shrivastava</td>
<td>205</td>
</tr>
<tr>
<td>Relationship Extraction based on Category of Medical Concepts from Lexical Contexts</td>
<td>Anupam Mondal, Dipankar Das and Sivaji Bandyopadhyay</td>
<td>212</td>
</tr>
<tr>
<td>Sinhala Word Joiner</td>
<td>Rajith Priyanga, Surangika Ranatunga and Gihan Dias</td>
<td>220</td>
</tr>
<tr>
<td>Supervised Methods For Ranking Relations In Web Search</td>
<td>Sumit Asthana and Asif Ekbal</td>
<td>227</td>
</tr>
<tr>
<td>Malayalam VerbFrames</td>
<td>Jisha P Jayan, Asha S Nair and Govindaru V</td>
<td>236</td>
</tr>
<tr>
<td>Hindi Shabdamitra: A Wordnet based E-Learning Tool for Language Learning and Teaching</td>
<td>Hanumant Redkar, Sandhya Singh, Dhara Gorasia, Meenakshi Somasundaram, Malhar Kulkarni and Pushpak Bhattacharyya</td>
<td>245</td>
</tr>
</tbody>
</table>
"A pessimist sees the difficulty in every opportunity; an optimist sees the opportunity in every difficulty"
– Understanding the psycho-sociological influences to it
Upendra Kumar, Vishal Kumar Rana, Srinivas Pykal and Amitava Das .......................... 255

End to End Dialog System for Telugu
Prathyusha Danda, Prathyusha Jwalapuram and Manish Shrivastava ........................... 265

Investigating how well contextual features are captured by bi-directional recurrent neural network models
Kushal Chawla, Sunil Kumar Sahu and Ashish Anand ................................. 273

Correcting General Purpose ASR Errors using Posteriors
Sunil Kumar Kopparapu and C. Anantaram ......................................................... 283

Retrieving Similar Lyrics for Music Recommendation System
Braja Gopal Patra, Dipankar Das and Sivaji Bandyopadhyay ................................. 290

Unsupervised Morpheme Segmentation Through Numerical Weighting and Thresholding
Joy Mahapatra and Sudip Kumar Naskar ......................................................... 298

Experiments with Domain Dependent Dialogue Act Classification using Open-Domain Dialogue Corpora
Swapnil Hingmire, Apoorv Shrivastava, Girish Palshikar and Saurabh Srivastava .......... 305

Normalization of Social Media Text using Deep Neural Networks
Ajay Shankar Tiwari and Sudip Kumar Naskar .................................................. 312

Aditya Thakker, Suhail Barot and Sudhir Bagul .................................................. 322

Exploring an Efficient Handwritten Manipuri Meetei-Mayek Character Recognition Using Gradient Feature Extractor and Cosine Distance Based Multiclass k-Nearest Neighbor Classifier
Kishorjit Nongmeikapam, Wahengbam Kumar and Mithlesh Prasad Singh .......... 328

A Modified Cosine-Similarity based Log Kernel for Support Vector Machines in the Domain of Text Classification
Rajendra Kumar Roul and Kushagr Arora .......................................................... 338

Document Embedding Generation for Cyber-Aggressive Comment Detection using Supervised Machine Learning Approach
Shylaja S S, Abhishek Narayanan, Abhijith Venugopal and Abhishek Prasad ........ 348

Coarticulatory propensity in Khalkha Mongolian
Ushashi Banerjee, Indranil Dutta and Irfan S ...................................................... 356

Developing Lexicon and Classifier for Personality Identification in Texts
Kumar Gourav Das and Dipankar Das .............................................................. 362

Linguistic approach based Transfer Learning for Sentiment Classification in Hindi
Vartika Rai, Sakshee Vijay and Dipti Misra .......................................................... 373
Scalable Bio-Molecular Event Extraction System towards Knowledge Acquisition
Pattabhi Rk Rao, Sindhuja Gopalan and Sobha Lalitha Devi .................................................. 383

Co-reference Resolution in Tamil Text
Vijay Sundar Ram and Sobha Lalitha Devi .......................................................... 392

Cross Linguistic Variations in Discourse Relations among Indian Languages
Sindhuja Gopalan, Lakshmi S and Sobha Lalitha Devi .................................................. 402

RULE BASED APPROACH OF CLAUSE BOUNDARY IDENTIFICATION IN TELUGU
Ganthoti Nagaraju, Thennarasu Sakkan and Christopher Mala ........................................ 408

Keynote Lecture 3: Towards Abstractive Summarization
Vasudeva Varma .......................................................... 417

"Who Mentions Whom?" - Understanding the Psycho-Sociological Aspects of Twitter Mention Network
R Sudhesh Solomon, Abhay Narayan, Srinivas P Y K L and Amitava Das ................................ 418

Study on Visual Word Recognition in Bangla across Different Reader Groups
Manjira Sinha and Tirthankar Dasgupta .......................................................... 427

Demystifying Topology of Autopilot Thoughts: A Computational Analysis of Linguistic Patterns of Psychological Aspects in Mental Health
Bibekananda Kundu and Sanjay Choudhury .......................................................... 435

A Deep Dive into Identification of Characters from Mahabharata
Apurba Paul and Dipankar Das .......................................................... 447

Neural Networks for Semantic Textual Similarity
Derek Prijatelj, Jugal Kalita and Jonathan Ventura .................................................. 456

Open Set Text Classification Using CNNs
Sridhama Prakhya, Vinodini Venkataram and Jugal Kalita ........................................ 466

Predicting User Competence from Linguistic Data
Yonas Woldemariam, Henrik Björklund and Suna Bensch ........................................ 476

Neural Morphological Disambiguation Using Surface and Contextual Morphological Awareness
Akhilesh Sudhakar and Anil Kumar Singh .......................................................... 485

Word Sense Disambiguation for Malayalam in a Conditional Random Field Framework
Junaida M K, Jisha P Jayan and Elizabeth Sherly .................................................. 495

Semisupervised Data Driven Word Sense Disambiguation for Resource-poor Languages
Pratibha Rani, Vikram Pudi and Dipti M. Sharma .................................................. 503

Notion of Semantics in Computer Science - A Systematic Literature Review
Sai Prasad Vrj Gollapudi and Venkatesh Choppella .................................................. 513
Tuesday, December 19, 2017

+ 9:00-9:30 Inaugural Ceremony

+ 9:30-10:30 Keynote Lecture 1 by Prof. Bjom W. Schuller

Keynote Lecture 1: NLP in Tomorrow’s Profiling - Words May Fail You
Björn W. Schuller

+ 10:30-11:00 Tea Break

+ 11:00-13:00 Technical Session I: Machine Translation and Speech:

Deriving Word Prosody from Orthography in Hindi
Somnath Roy

Three-phase training to address data sparsity in Neural Machine Translation
Ruchit Agrawal, Mihir Shekhar and Dipti Sharma

Reference Scope Identification for Citances Using Convolutional Neural Networks
Saurav Jha, Aanchal Chaurasia, Akhilesh Sudhakar and Anil Kumar Singh

A vis-a-vis evaluation of MT paradigms for linguistically distant languages
Ruchit Agrawal, Jahfar Ali and Dipti Misra Sharma

+ 11:00-13:00 Technical Session II : Text Categorization:

Textual Relations and Topic-Projection: Issues in Text Categorization
Lahari Chatterjee, Samir Karmakar and Abahan Datta

POS Tagging For Resource Poor Languages Through Feature Projection
Pruthwik Mishra, Vandan Mujadia and Dipti Misra Sharma

An Exploration of Word Embedding Initialization in Deep-Learning Tasks
Tom Kocmi and Ondrej Bojar
Technical Session III: Parsing Code-mixed Data:

Curriculum Design for Code-switching: Experiments with Language Identification and Language Modeling with Deep Neural Networks
Monojit Choudhury, Kalika Bali, Sunayana Sitaram and Ashutosh Baheti

Quantitative Characterization of Code Switching Patterns in Complex Multi-Party Conversations: A Case Study on Hindi Movie Scripts
Adithya Pratapa and Monojit Choudhury

Towards Normalising Konkani-English Code-Mixed Social Media Text
Akshata Phadte and Gaurish Thakkar

Towards developing a phonetically balanced code-mixed speech corpus for Hindi-English ASR
Ayushi Pandey, Brij Mohan Lal Srivastava and Suryakanth Gangashetty

Lunch

Keynote Lecture 2 by Prof. NG Hwee Tou

Keynote Lecture 2: Grammatical Error Correction: Past, Present and Future
NG Hwee Tou

Technical Session IV: Information Extraction:

Hybrid Approach for Marathi Named Entity Recognition
Nita Patil, Ajay Patil and B.V. Pawar

Sentiment Analysis: An Empirical Comparative Study of Various Machine Learning Approaches
Swapnil Jain, Shrikant Malviya, Rohit Mishra and Uma Shanker Tiwary
+ 15:00-16:30 Technical Session V : Discourse and Dialogue:

Handling Multi-Sentence Queries in a Domain Independent Dialogue System
Prathyusha Jwalapuram and Radhika Mamidi

Document Level Novelty Detection: Textual Entailment Lends a Helping Hand
Tanik Saikh, Tirthankar Ghosal, Asif Ekbal and Pushpak Bhattacharyya

Is your Statement Purposeless? Predicting Computer Science Graduation Admission Acceptance based on Statement Of Purpose
Diptesh Kanojia, Nikhil Wani and Pushpak Bhattacharyya

+ 15:00-16:30 Technical Session VI : Lexical Analysis:

Natural Language Programming with Automatic Code Generation towards Solving Addition-Subtraction Word Problems
Sourav Mandal and Sudip Kumar Naskar

Unsupervised Separation of Transliterable and Native Words for Malayalam
Deepak P

Known Strangers: Cross Linguistic Patterns in Multilingual Multidirectional Dictionaries
Rejitha K. S. and Rajesha N.

+ 16:30-17:30 Tea Break

+ 16:30-17:30 Poster and Demo Session-I:

Tutorial for Deaf Teaching Punjabi Alphabet using Synthetic Animations
Lalit Goyal and Vishal Goyal

SemTagger: A Novel Approach for Semantic Similarity Based Hashtag Recommendation on Twitter
Kuntal Dey, Ritvik Shrivastava, Saroj Kaushik and L. Venkata Subramaniam

Reasoning with Sets to Solve Simple Word Problems Automatically
Sowmya S Sundaram and Deepak Khemani

Improving NER for Clinical Texts by Ensemble Approach using Segment Representations
Hamada Nayel and H L Shashirekha
Beyond Word2Vec: Embedding Words and Phrases in Same Vector Space
Vijay Prakash Dwivedi and Manish Shrivastava

Relationship Extraction based on Category of Medical Concepts from Lexical Contexts
Anupam Mondal, Dipankar Das and Sivaji Bandyopadhyay

Sinhala Word Joiner
Rajith Priyanga, Surangika Ranatunga and Gihan Dias

Supervised Methods For Ranking Relations In Web Search
Sumit Asthana and Asif Ekbal

Malayalam VerbFrames
Jisha P Jayan, Asha S Nair and Govindaru V

Hindi Shabdamitra: A Wordnet based E-Learning Tool for Language Learning and Teaching
Hanumant Redkar, Sandhya Singh, Dhara Gorasia, Meenakshi Somasundaram, Malhar Kulkarni and Pushpak Bhattacharyya

"A pessimist sees the difficulty in every opportunity; an optimist sees the opportunity in every difficulty" – Understanding the psycho-sociological influences to it
Updendra Kumar, Vishal Kumar Rana, Srinivas Pykl and Amitava Das

End to End Dialog System for Telugu
Prathyusha Danda, Prathyusha Jwalapuram and Manish Shrivastava

Investigating how well contextual features are captured by bi-directional recurrent neural network models
Kushal Chawla, Sunil Kumar Sahu and Ashish Anand

Correcting General Purpose ASR Errors using Posteriors
Sunil Kumar Kopparapu and C. Anantaram

Retrieving Similar Lyrics for Music Recommendation System
Braja Gopal Patra, Dipankar Das and Sivaji Bandyopadhyay

Unsupervised Morpheme Segmentation Through Numerical Weighting and Thresholding
Joy Mahapatra and Sudip Kumar Naskar
Experiments with Domain Dependent Dialogue Act Classification using Open-Domain Dialogue Corpora
Swapnil Hingmire, Apoorv Shrivastava, Girish Palshikar and Saurabh Srivastava

Normalization of Social Media Text using Deep Neural Networks
Ajay Shankar Tiwari and Sudip Kumar Naskar

Aditya Thakker, Suhail Barot and Sudhir Bagul

Exploring an Efficient Handwritten Manipuri Meetei-Mayek Character Recognition Using Gradient Feature Extractor and Cosine Distance Based Multiclass k-Nearest Neighbor Classifier
Kishorjit Nongmeikapam, Wahengbam Kumar and Mithlesh Prasad Singh

A Modified Cosine-Similarity based Log Kernel for Support Vector Machines in the Domain of Text Classification
Rajendra Kumar Roul and Kushagr Arora

Document Embedding Generation for Cyber-Aggressive Comment Detection using Supervised Machine Learning Approach
Shylaja S S, Abhishek Narayanan, Abhijith Venugopal and Abhishek Prasad

Coarticulatory propensity in Khalkha Mongolian
Ushashi Banerjee, Indranil Dutta and Irfan S

Developing Lexicon and Classifier for Personality Identification in Texts
Kumar Gourav Das and Dipankar Das

Linguistic approach based Transfer Learning for Sentiment Classification in Hindi
Vartika Rai, Sakshee Vijay and Dipti Misra

Scalable Bio-Molecular Event Extraction System towards Knowledge Acquisition
Pattabhi Rk Rao, Sindhuja Gopalan and Sobha Lalitha Devi

Co-reference Resolution in Tamil Text
Vijay Sundar Ram and Sobha Lalitha Devi

Cross Linguistic Variations in Discourse Relations among Indian Languages
Sindhuja Gopalan, Lakshmi S and Sobha Lalitha Devi
Tuesday, December 19, 2017 (continued)

**RULE BASED APPROACH OF CLAUSE BOUNDARY IDENTIFICATION IN TELUGU**
Ganthoti Nagaraju, Thennarasu Sakkan and Christopher Mala

+ 17:30-19:30 NLPAI Meeting
+ 19:00-20:00 Cultural Programme
+ 20:00-Onwards Dinner

Wednesday, December 20, 2017

+ 9:30-10:30 Keynote Lecture 3 by Vasudeva Varmaa

*Keynote Lecture 3: Towards Abstractive Summarization*
Vasudeva Varma

+ 10:30-11:00 Tea Break

+ 11:00-13:00 Technical Session VII: Socio-Psycho Text Analysis: Emerging Trends

"Who Mentions Whom?" - *Understanding the Psycho-Sociological Aspects of Twitter Mention Network*
R Sudhesh Solomon, Abhay Narayan, Srinivas P Y K L and Amitava Das

*Study on Visual Word Recognition in Bangla across Different Reader Groups*
Manjira Sinha and Tirthankar Dasgupta

*Demystifying Topology of Autopilot Thoughts: A Computational Analysis of Linguistic Patterns of Psychological Aspects in Mental Health*
Bibekananda Kundu and Sanjay Choudhury

*A Deep Dive into Identification of Characters from Mahabharata*
Apurba Paul and Dipankar Das
Wednesday, December 20, 2017 (continued)

+ 11:00-13:00 Technical Session VIII: Deep Neural Networks:

*Neural Networks for Semantic Textual Similarity*
Derek Prijatelj, Jugal Kalita and Jonathan Ventura

*Open Set Text Classification Using CNNs*
Sridhama Prakhya, Vinodini Venkataram and Jugal Kalita

*Predicting User Competence from Linguistic Data*
Yonas Woldemariam, Henrik Björklund and Suna Bensch

*Neural Morphological Disambiguation Using Surface and Contextual Morphological Awareness*
Akhilesh Sudhakar and Anil Kumar Singh

+ 11:00-13:00 Technical Session IX: Semantics:

*Word Sense Disambiguation for Malayalam in a Conditional Random Field Framework*
Junaida M K, Jisha P Jayan and Elizabeth Sherly

*Semisupervised Data Driven Word Sense Disambiguation for Resource-poor Languages*
Pratibha Rani, Vikram Pudi and Dipti M. Sharma

*Notion of Semantics in Computer Science - A Systematic Literature Review*
Sai Prasad Vrj Gollapudi and Venkatesh Choppella

*Semantic Enrichment Across Language: A Case Study of Czech Bibliographic Databases*
Pavel Smrz and Lubomir Otrusina
Wednesday, December 20, 2017 (continued)

+ 13:00-14:00 Lunch Break
+ 14:00-15:00 Industry Talk
+ 15:00-15:30 Tea Break
+ 15:30-17:00 Technical Session X: Student Paper Contest
+ 15:30-17:00 Technical Session XI: NLP Tools Contest
+ 17:00-17:30 Valedictory Session