

Table of Contents – Part II

Sentiment Analysis and Emotion Recognition

| | |
|--|----|
| Sentence-Level Sentiment Analysis in the Presence of Modalities (Invited Paper) | 1 |
| <i>Yang Liu, Xiaohui Yu, Bing Liu, and Zhongshuai Chen</i> | |
| Word-Level Emotion Recognition Using High-Level Features (Invited Paper) | 17 |
| <i>Johanna D. Moore, Leimin Tian, and Catherine Lai</i> | |
| Constructing Context-Aware Sentiment Lexicons with an Asynchronous Game with a Purpose | 32 |
| <i>Marina Boia, Claudiu Cristian Musat, and Boi Faltings</i> | |
| Acknowledging Discourse Function for Sentiment Analysis | 45 |
| <i>Phillip Smith and Mark Lee</i> | |
| A Method of Polarity Computation of Chinese Sentiment Words Based on Gaussian Distribution | 53 |
| <i>Ruijing Li, Shumin Shi, Heyan Huang, Chao Su, and Tianhang Wang</i> | |
| A Sentence Vector Based Over-Sampling Method for Imbalanced Emotion Classification | 62 |
| <i>Tao Chen, Ruifeng Xu, Qin Lu, Bin Liu, Jun Xu, Lin Yao, and Zhenyu He</i> | |
| Reader Emotion Prediction Using Concept and Concept Sequence Features in News Headlines | 73 |
| <i>Yuanlin Yao, Ruifeng Xu, Qin Lu, Bin Liu, Jun Xu, Chengtian Zou, Li Yuan, Shuwei Wang, Lin Yao, and Zhenyu He</i> | |
| Identifying the Targets of the Emotions Expressed in Health Forums . . . | 85 |
| <i>Sandra Bringay, Eric Kergosien, Pierre Pompidor, and Pascal Poncelet</i> | |
| Investigating the Role of Emotion-Based Features in Author Gender Classification of Text | 98 |
| <i>Calkin Suero Montero, Myriam Munezero, and Tuomo Kakkonen</i> | |

Opinion Mining and Social Networks

| | |
|---|-----|
| A Review Corpus for Argumentation Analysis | 115 |
| <i>Henning Wachsmuth, Martin Trenkmann, Benno Stein, Gregor Engels, and Tsvetomira Palakarska</i> | |

| | |
|--|-----|
| Looking for Opinion in Land-Use Planning Corpora | 128 |
| <i>Eric Kergosien, Cédric Lopez, Mathieu Roche, and Maquelonne Teisseire</i> | |
| Cross-Lingual Product Recommendation Using Collaborative Filtering with Translation Pairs | 141 |
| <i>Kanako Komiya, Shohei Shibata, and Yoshiyuki Kotani</i> | |
| Identifying a Demand Towards a Company in Consumer-Generated Media | 153 |
| <i>Yuta Kikuchi, Hiroya Takamura, Manabu Okumura, and Satoshi Nakazawa</i> | |
| Standardizing Tweets with Character-Level Machine Translation | 164 |
| <i>Nikola Ljubešić, Tomaž Erjavec, and Darja Fišer</i> | |
| #impressme: The Language of Motivation in User Generated Content | 176 |
| <i>Marc T. Tomlinson, David B. Bracewell, Wayne Krug, and David Hinote</i> | |
| Mining the Personal Interests of Microbloggers via Exploiting Wikipedia Knowledge | 188 |
| <i>Miao Fan, Qiang Zhou, and Thomas Fang Zheng</i> | |
| Website Community Mining from Query Logs with Two-Phase Clustering | 201 |
| <i>Lidong Bing, Wai Lam, Shoaib Jameel, and Chunliang Lu</i> | |
| Extracting Social Events Based on Timeline and User Reliability Analysis on Twitter | 213 |
| <i>Bayar Tsolmon and Kyung-Soon Lee</i> | |

Machine Translation and Multilingualism

| | |
|--|-----|
| Beam-Width Adaptation for Hierarchical Phrase-Based Translation | 224 |
| <i>Fei Su, Gang Chen, Xinyan Xiao, and Kaile Su</i> | |
| Training Phrase-Based SMT without Explicit Word Alignment | 233 |
| <i>Cyrine Nasri, Kamel Smaili, and Chiraz Latiri</i> | |
| Role of Paraphrases in PB-SMT | 242 |
| <i>Santanu Pal, Pintu Lohar, and Sudip Kumar Naskar</i> | |
| Inferring Paraphrases for a Highly Inflected Language from a Monolingual Corpus | 254 |
| <i>Kfir Bar and Nachum Dershowitz</i> | |

| | |
|---|-----|
| Improving Egyptian-to-English SMT by Mapping Egyptian into MSA | 271 |
| <i>Nadir Durrani, Yaser Al-Onaizan, and Abraham Ittycheriah</i> | |
| Bilingually Learning Word Senses for Translation | 283 |
| <i>João Casteleiro, Gabriel Pereira Lopes, and Joaquim Silva</i> | |
| Iterative Bilingual Lexicon Extraction from Comparable Corpora with Topical and Contextual Knowledge (Best Student Paper Award) | 296 |
| <i>Chenhui Chu, Toshiaki Nakazawa, and Sadao Kurohashi</i> | |
| Improving Bilingual Lexicon Extraction from Comparable Corpora Using Window-Based and Syntax-Based Models | 310 |
| <i>Amir Hazem and Emmanuel Morin</i> | |
| An IR-Based Strategy for Supporting Chinese-Portuguese Translation Services in Off-line Mode..... | 324 |
| <i>Jordi Centelles, Marta Ruiz Costa-jussà, Rafael E. Banchs, and Alexander Gelbukh</i> | |
| Cross Lingual Snippet Generation Using Snippet Translation System ... | 331 |
| <i>Pintu Lohar, Pinaki Bhaskar, Santanu Pal, and Sivaji Bandyopadhyay</i> | |
| A Novel Machine Translation Method for Learning Chinese as a Foreign Language | 343 |
| <i>Tiansi Dong and Armin B. Cremers</i> | |

Information Retrieval

| | |
|---|-----|
| A New Relevance Feedback Algorithm Based on Vector Space Basis Change | 355 |
| <i>Rabeb Mbarek, Mohamed Tmar, and Hawete Hattab</i> | |
| How Complementary Are Different Information Retrieval Techniques? A Study in Biomedicine Domain | 367 |
| <i>Xiangdong An and Nick Cercone</i> | |
| Performance of Turkish Information Retrieval: Evaluating the Impact of Linguistic Parameters and Compound Nouns | 381 |
| <i>Hatem Haddad and Chedi Bechikh Ali</i> | |

Text Classification and Clustering

| | |
|--|-----|
| How Document Properties Affect Document Relatedness Measures (Verifiability Award) | 392 |
| <i>Jessica Perrie, Aminul Islam, and Evangelos Milios</i> | |

| | |
|--|-----|
| Multi-attribute Classification of Text Documents as a Tool for Ranking and Categorization of Educational Innovation Projects | 404 |
| <i>Alexey An, Bakytkan Dauletbakov, and Eugene Levner</i> | |
| Named Entities as New Features for Czech Document Classification | 417 |
| <i>Pavel Král</i> | |
| A Knowledge-Poor Approach to Turkish Text Categorization | 428 |
| <i>Savaş Yıldırım</i> | |
| Credible or Incredible? Dissecting Urban Legends | 441 |
| <i>Marco Guerini and Carlo Strapparava</i> | |
| Intelligent Clustering Scheme for Log Data Streams | 454 |
| <i>Basanta Joshi, Umanga Bista, and Manoj Ghimire</i> | |

Text Summarization

| | |
|--|-----|
| Graph Ranking on Maximal Frequent Sequences for Single Extractive Text Summarization | 466 |
| <i>Yulia Ledeneva, René Arnulfo García-Hernández, and Alexander Gelbukh</i> | |

Plagiarism Detection

| | |
|---|-----|
| A Graph Based Automatic Plagiarism Detection Technique to Handle Artificial Word Reordering and Paraphrasing (Best Paper Award) | 481 |
| <i>Niraj Kumar</i> | |
| Identification of Plagiarism Using Syntactic and Semantic Filters | 495 |
| <i>R. Vijay Sundar Ram, Efsthios Stamatatos, and Sobha Lalitha Devi</i> | |

Style and Spelling Checking

| | |
|--|-----|
| Readability Classification of Bangla Texts | 507 |
| <i>Zahurul Islam, Md. Rashedur Rahman, and Alexander Mehler</i> | |
| State-of-the-Art in Weighted Finite-State Spell-Checking | 519 |
| <i>Tommi A. Pirinen and Krister Lindén</i> | |
| Spelling Correction for Kazakh | 533 |
| <i>Aibek Makazhanov, Olzhas Makhambetov, Islam Sabyrgaliyev, and Zhandos Yessenbayev</i> | |

Speech Processing

| | |
|--|-----|
| A Preliminary Study on the VOT Patterns of the Assamese Language and Its Nalbaria Variety | 542 |
| <i>Sanghamitra Nath, Himangshu Sarma, and Utpal Sharma</i> | |

Applications

| | |
|---|-----|
| Evaluation of Sentence Compression Techniques Against Human Performance..... | 553 |
| <i>Prasad Perera and Leila Kosseim</i> | |

| | |
|---|-----|
| Automatically Assessing Children’s Writing Skills Based on Age-Supervised Datasets | 566 |
| <i>Nelly Moreno, Sergio Jimenez, and Julia Baquero</i> | |

| | |
|---------------------------|-----|
| Author Index | 579 |
|---------------------------|-----|