Table of Contents

Improving Summarization Performance by Sentence Compression — A Pilot Study
Chin-Yew Lin 1

A Practical Text Summarizer by Paragraph Extraction for Thai
Chuleerat Jaruskulchai and Canasai Kruengkrai 9

An Approach for Combining Content-based and Collaborative Filters
Qing Li and Byeong Man Kim 17

A Differential LSI Method for Document Classification
Liang Chen, Naoyuki Tokuda and Akira Nagai 25

Poisson Naive Bayes for Text Classification with Feature Weighting
Sang-Bum Kim, Hee-Cheol Seo and Hae-Chang Rim 33

Text Classification in Asian Languages without Word Segmentation
Fuchun Peng, Xiangji Huang, Dale Schuurmans and Shaojun Wang 41

Feature Selection in Categorizing Procedural Expressions
Mineki Takechi, Takenobu Tokunaga, Yuji Matsumoto and Hozumi Tanaka 49

Learning Bilingual Translations from Comparable Corpora to Cross-Language Information Retrieval: Hybrid Statistics-based and Linguistics-based Approach
Fatiha Sadat, Masatoshi Yoshikawa and Shunsuke Uemura 57

BRIDJE over a Language Barrier: Cross-Language Information Access by Integrating Translation and Retrieval
Tetsuya Sakai, Makoto Koyama, Masaru Suzuki, Akira Kumano and Toshihiko Manabe 65

Issues in Pre- and Post-translation Document Expansion: Untranslatable Cognates and Missegmented Words
Gina-Anne Levow 77

Very Low Dimensional Latent Semantic Indexing for Local Query Regions
Yinghui Xu and Kyoji Umemura 84

AnyQ: Answer Set based Information Retrieval System
Hyo-Jung Oh, Myung-Gil Jang and Moon-Soo Chang 92

Dynamic Programming Matching for Large Scale Information Retrieval
Eiko Yamamoto, Masahiro Kishida, Yoshinori Takenami, Yoshiyuki Takeda and Kyoji Umemura 100

Improving Document Clustering by Utilizing Meta-Data
Kam-Fai Wong, Nam-Kiu Chan and Kam-Lai Wong 109

Temporal Ranking for Fresh Information Retrieval
Nobuyoshi Sato, Minoru Uehara and Yoshifumi Sakai 116

Extraction of User Preferences from a Few Positive Documents
Byeong Man Kim, Qing Li and Jong Wan Kim 124

Keyword-based Document Clustering
Seung-Shik Kang 132

Text Categorization Using Automatically Acquired Domain Ontology
Shih-Hung Wu, Tzong-Han Tsai and Wen-Lian Hsu 138

A Sentence Reduction using Syntax Control
Minh Le Nguyen and Susumu Horiguchi 146

Cross-Language Information Retrieval Based on Category Matching Between Language Versions of a Web Directory
Fuminori Kimura, Akira Maeda, Masatoshi Yoshikawa and Uemura Shunsuke 153

Korean Named Entity Recognition using HMM and CoTraining Model
Euisok Chung, Yi-Gyu Hwang and Myung-Gil Jang 161